EVALUATION AND IMPLEMENTATION OF PROXIMAL AND REMOTE SENSING TECHNIQUES IN A SORGHUM BREEDING PROGRAM

A Dissertation

by

NICHOLAS ACE PUGH

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	William L. Rooney
Committee Members,	Cristine L. S. Morgan
	Nithya Rajan
	Jinha Jung
Head of Department,	David D. Baltensperger

December 2018

Major Subject: Plant Breeding

ABSTRACT

Plant breeders must improve the efficiency of their programs to meet the needs of a rising world population and a rapidly changing climate. Proximal and remote sensing technologies stand as an attractive option for breeders, as they can help to alleviate the phenotyping bottleneck in most programs. Sorghum (*Sorghum bicolor*, L. Moench) stands as an excellent crop for researchers to test these technologies. Traits such as biomass yield, plant height, and anthracnose resistance are all important traits to sorghum breeders. In addition, spatial variability is often of chief concern in sorghum breeding trials. For biomass yield and plant height, we utilized Canopy Volume/Canopy Coverage estimates and structure-from-motion (SfM)-derived plant height estimates, respectively. It was determined that Canopy Volume is effective at teasing out the most genotypic variation, while minimizing error, when compared to Canopy Coverage as an estimate of biomass. Additionally, the ninety-fifth percentile (P95) of SfM height estimate was shown to be more strongly correlated with ground-truth estimates of plant height. For estimation of spatial variability, we utilized soil apparent electrical conductivity (EC_a) grids to test three different linear models for their ability to account for spatial autocorrelation in various trials via a Moran's I test. We found that in most situations it is unnecessary to use blocking or soil EC_a grids, but that in situations where it is necessary there is usually a superior choice. Thus, it is probably best to use both methods and determine which is the best for each situation by finding which model removes spatial autocorrelation. Finally, we tested normalized difference vegetation index (NDVI) for its ability to estimate the incidence and severity of anthracnose (*Colletotrichum sublineola*) disease in a sorghum breeding population. We find that NDVI is highly effective at delineating genotypic variation for anthracnose resistance later in the season but is inferior to ground-truth methods early on. There are many other traits and aspects of

sorghum breeding that have yet to be tested, but these examples serve as an indicator that remote and proximal sensing will be an incredibly valuable tool in the future.

DEDICATION

For Mom and Dad, who have always supported me throughout my graduate career

ACKNOWLEDGEMENTS

I would like to thank my colleagues and friends in the Sorghum Breeding Program at Texas A & M including especially for their support on these projects, and for helping to make my experience as a doctoral student an enjoyable one. I would also like to thank the Crop Testing Program for their contributions to the second project in this dissertation, and the Corpus Christi group as well as Dr. Lonesome Malambo, Dr. Xiongzhe Han, Dr. Sorin Popescu, and Dr. J. Alex Thomasson for their contributions to the first and third project. I would also like to thank all the students, staff, and faculty at Texas A & M University, including David Horne, Jacob Pekar, and Steven Anderson, especially, in the Soil and Crop Sciences department, that helped me to learn the materials and gain the skills that were required to finish this document. Particularly, I would like to thank Dr. Seth Murray and his program for initially sparking my interest in plant breeding as a discipline and for allowing me to work as an intern in his program.

I would like to thank my committee, Dr. Cristine Morgan, Dr. Jinha Jung, and Dr. Nithya Rajan, for their contributions and support over the course of this study. They have helped me to learn to critically examine my work and have offered a great deal of insight about how to improve the studies contained herein. I have enjoyed my time collaborating with all of them and look forward to maintaining a connection with them after the conclusion of this work.

I want to thank my family as well, including my Mom, Deborah Pugh, my Dad, Randall Pugh, my sister, Brittany Pugh, and my sister's amazing dog, Louie. You all helped to support me and get me through the tougher days of being a Ph.D. student, and I will never forget what you have done for me. To Brittany, I thank you immensely for supporting me as a student and for letting me stay at your apartment on some of my trips to collect data or present at conferences. I

had a good time on these trips since I got to hang out with you and Louie. Mom and Dad, I would not even be in graduate school if you had not supported me and helped me to get my undergraduate education. I have enjoyed having the opportunity to teach aspects of my job to you, but I have also continued to learn new things from you even as a graduate student.

I would be remiss if I did not also thank my numerous friends outside of my department that helped me grow and helped me survive the trials and tribulations of graduate school. These fellow graduate students helped me to have fun along the way, and I also learned a great deal from them about their respective fields and, I hope, taught them about mine as well. Though this list is not exhaustive, I want to personally thank Reed Stubbendieck, Caitlin Curry, David Forgacs, Ashley Mattison, Drew Anderson, David Green, Michael Werry, and Scott Mattison for the good times that I have had with them during my time here. I will continue to maintain connections with these individuals and will continue to consider them the truest friends. Similarly, I want to thank my best friends in Oklahoma City that I have known since middle school or even elementary school, including Christopher Frazier, Timothy Fadum, Michael Fadum, Shawn Fadum, and Daniel Fawcett. All of you helped me to unwind and take a breather when I was able to visit home during the holidays.

Finally, I would like to thank my committee chair and advisor, Dr. William Rooney, for his instruction and guidance over the course of my graduate career. I have learned a great deal about plant breeding and, perhaps more importantly, a great deal about being a good scientist and a good leader. I have learned to improve my skills as a writer, while simultaneously learning how to manage and conduct field work in a large breeding program. I have come a long way under his instruction compared to where I originally started, and I look forward to taking the knowledge he has given me and using it to build a successful career.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Professor William L. Rooney, Professor Cristine L. S. Morgan, and Professor Nithya Rajan of the Soil and Crop Sciences Department at Texas A & M University, College Station, TX as well as Professor Jinha Jung at Texas A & M Corpus Christi. Financial support for these projects was provided by Texas A & M Agrilife, by the DOE-ARPA-E TERRA program grant no. DE-AR0000596, and by the United Sorghum Checkoff Program.

For the first study, rotary-winged UAS flights were conducted by Dr. Lonesome Malambo and the Corpus Christi Flight Team in College Station and Corpus Christi, respectively. Analysis of the raw imagery was conducted by Dr. Lonesome Malambo and Dr. Anjin Chang with guidance by Dr. Sorin Popescu and Dr. Jinha Jung, respectively. Ground-truth estimates of biomass yield and plant height were obtained by the student as well as by the Sorghum Breeding Program at Texas A & M. Analysis of the biomass yield and plant height estimates derived from the imagery was conducted by Mr. Ace Pugh independently with guidance from Dr. William L. Rooney

For the second study, instruction on the use of the EM38-MK2 was given by Ms. Sarah Vaughan and Ms. Catherine Kobylinksi in Dr. Cristine Morgan's program. Access to the trial locations and agronomic data for the sorghum performance trials was granted by Mr. Dennis Pietsch and Ms. Katrina Horn in the Crop Testing Program. Acquisition of soil ECa data, construction of ECa grids, and subsequent statistical analysis was conducted independently by Mr. Ace Pugh with guidance from Dr. Cristine Morgan and Dr. William Rooney.

For the third and final study, preparation and application of the anthracnose inoculant was conducted by Dr. Thomas S. Isakeit and the Texas A & M Sorghum Breeding Program. Fixed-wing UAS flights were conducted by Dr. Dale Cope and his flight team in College Station, TX. Analysis of the raw imagery was conducted by Dr. Anjin Chang and Dr. Jinha Jung at Texas A & M Corpus Christi, and derivation of NDVI estimates from the orthorectified imagery was conducted by Dr. Xiongzhe Han in Dr. J. Alex Thomasson's program. Ground-truth estimates of anthracnose incidence and severity were recorded by Mr. Delroy Collins in the Texas A & M Sorghum Breeding Program. Analysis of the NDVI and ground-truth data was conducted by Mr. Ace Pugh independently with the guidance of Dr. William Rooney.

NOMENCLATURE

AUDPC	Area Under the Disease Progress Curve
BMYLD	Biomass Yield
CC 16	College Station 2016
CHM	Canopy Height Model
COV	Canopy Cover
CRD	Completely Randomized Design
CS 17	College Station 2017
DAP	Days After Planting
DEM	Digital Elevation Map
DSM	Digital Surface Map
EC_a	Apparent Electrical Conductivity
EMI	Electromagnetic Induction
GT	Ground-truth
MAX	Maximum Percentile of Structure-from-motion Data
NDVI	Normalized Difference Vegetation Index
P95	Ninety-fifth Percentile of Structure-from-motion Data
PI	Photoperiod Insensitive
PS	Photoperiod Sensitive
RCBD	Randomized Complete Block Design
REML	Restricted Maximum Likelihood
SfM	Structure-from-motion
TAMU	Texas A&M University

UAS	Unmanned Aerial System
VOL	Canopy Volume

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION AND LITERATURE REVIEW

1.1 The Need for Proximal and Remote Sensing in Plant Breeding

Our world faces numerous challenges in the coming decades. Two of the challenges that are of utmost concern to crop improvement scientists are global climate change exacerbated by the activities of mankind and a rapid expected population increase (Godfray et al., 2010; Cairns et al., 2013; Wheeler and von Braun, 2013). The rate of increase in crop yields that is necessary to address these demands is currently not being achieved (Ray et al., 2012; Ray et al., 2013). Monetary and time-oriented costs have led modern breeding programs to an impasse in their phenotyping capability. Simply put, phenotyping has become the primary bottleneck and the main obstacle when breeding programs are attempting to increase their rate of genetic gain (Furbank and Tester, 2011). To rise to this challenge, high-throughput phenotyping techniques, especially those using sensors mounted on aerial mobile platforms, are becoming increasingly enticing to researchers (Tester and Langridge, 2010; Araus and Cairns, 2014; Sankaran et al., 2015; Shi et al., 2016).

Though plant breeders traditionally are most concerned with food and fiber yields when they are selecting genotypes to advance within their programs, there are many facets of these traits that can be considered (Fehr, 1987). Additionally, there are often other phenotypes that are of vast importance to breeders, such as plant height, disease and pest resistance, and drought tolerance. However, while there are many possible options available to crop scientists and understanding them all can seem daunting because of their relatively recent appearance. Among these are several

proximal and remote sensing technologies that can be implemented into existing programs.

High-throughput phenotyping platforms outfitted with a variety of sensors are becoming more enticing to crop improvement scientists. Platforms that can be outfitted with sensor payloads include but are not limited to ground-based vehicles such as tractors, satellites, as well as fixed-wing or rotary-wing unmanned aerial systems or UASs. Perhaps the most frequently used and tested as of late are unmanned aerial systems, which have become very popular as subjects for high-throughput phenotyping studies due to their low cost and efficiency (Chapman et al., 2014; Sankaran et al., 2015; Shi et al., 2016; Yue et al., 2017). While satellite imagery has also been utilized in the past as well as in present studies to acquire vast amounts of information over a very large area, UASs provide certain advantages over more traditional methods in that they are less affected by weather, more flexible in terms of logistics such as scheduling, and their flight paths can be easily adjusted per changing conditions in the field (Stroppiana et al., 2015).

UAS can be equipped with several possible sensors including RGB/CIR (red, green, blue / color-infrared) cameras, LiDAR (light detection and ranging), multispectral cameras, hyperspectral imagers, long-wave infrared cameras or thermal imaging cameras, as well as conventional digital cameras (Araus and Cairns, 2014). One of the caveats of UASs being used for high-throughput phenotyping applications has traditionally been their limited overall payload, being unable to carry as much as ground-based systems (Stroppiana et al., 2015). This causes most researchers to use a few

standard or modified cameras available commercially (Lelong et al., 2008; Hunt et al., 2010; Stroppiana et al., 2015). Despite these caveats, the sheer amount of data that can be generated by remote sensing platforms could conceivably outweigh any possible shortcomings when compared to traditional measurements of phenotypic characteristics taken manually by researchers in the field. Perhaps one of the greatest challenges facing breeders in an age of high-throughput phenotyping will be how to finesse the incredibly large datasets that will be generated.

1.2 Estimation of Biomass Yield and Plant Height in Sorghum Using Unmanned Aerial Systems

Breeding crops for biomass productivity has been important even before the adoption of scientific approaches to breeding began to develop over a century ago. Today, breeding for biomass productivity takes many forms; in sugarcane, sugar and bioenergy production are priorities, and in alfalfa, yield, quality, and stand persistence are important. Regardless of the biomass crop, biomass yield is a common and critically important trait. Plant breeders are constantly working towards improving biomass yield so that they are a more attractive forage, lumber or alternative for energy producers. However, alongside other notable factors, phenotyping has heretofore remained one of the primary bottlenecks in modern breeding programs (Furbank and Tester, 2011). Therefore, high-throughput methodologies must be developed to ensure that bioenergy crops can be efficiently improved (Walter et al., 2012; Araus and Cairns, 2014).

As mentioned, there are several different crops that are grown for their superior biomass properties. Perhaps one of the most promising of these is sorghum (*Sorghum*

bicolor L. Moench) (Rooney et al., 2007; Calviño, 2012). At non-Equatorial latitudes, photoperiod-sensitive sorghum maintains vegetative growth for a much longer time than photoperiod-insensitive genotypes resulting in higher biomass production (Rooney and Aydin, 1999; Murphy et al., 2011). Of the many traits of interest, perhaps the most important are plant height and biomass yield, wherein the former is often an enormously important component of the latter (Machado et al., 2002; Salas Fernandez et al., 2009; Yin et al., 2011; Roth and Streit, 2018). Thus, breeding for the improvement of both forage and bioenergy sorghum is quite different from the approach that would be followed when breeding for increased grain yields in the same crop. Additionally, the traits mentioned can be quite laborious to phenotype, thereby creating the impetus for the evaluation of high-throughput techniques (Furbank and Tester, 2011; White et al., 2012).

One potential method that breeders can use to increase the efficiency and accuracy of their efforts to improve their crops is to use remote sensing via unmanned aerial systems (UAS) and photogrammetry (Colomina and Molina, 2014; Shi et al., 2016; Yue et al., 2017; Roth and Streit, 2018; Shafian et al., 2018). Several recent studies have already demonstrated that photogrammetry can potentially be used to predict biomass or height in several crops including maize (*Zea mays*), wheat (*Triticum aestivum*), sorghum, sugarcane (*Saccharum officinarum*), barley (*Hordeum vulgare*), soybean (*Glycine max*), poppy (*Papaver somniferum*), and grapes (*Vitis vinifera*) (Grenzdörffer,; Bendig et al., 2015; Holman et al., 2016; Shi et al., 2016; Hämmerle and Höfle, 2016; Watanabe et al., 2017; De Souza et al., 2017; Matese et al., 2017; Iqbal et

al., 2017; Yue et al., 2017; Chang et al., 2017; Pugh et al., 2017; Roth and Streit, 2018; Malambo et al., 2018). It is conceivable that sorghum scientists could utilize UAS to predict total biomass yield in sorghum, and temporal monitoring of the total growth and development of canopy height and volume could have incredible potential benefits (Cooper et al., 2016; Pugh et al., 2017).

Researchers that use UAV methodology could also elucidate details that were previously difficult or impossible to describe using traditional techniques. For example, constructing high-resolution multi-temporal growth curves is a novel use for UAS technology, wherein plant breeders could investigate the overall vigor of a genotype over the entire growing period (Cooper et al., 2016; Pugh et al., 2017). Additional applications could include delineating maturity subgroups within breeding populations, determining optimum timepoints for which to harvest each genotype, and assessing overall response to stresses (Cooper et al., 2016; Pugh et al., 2017).

1.3 Proximal Sensing of the Apparent Electrical Conductivity of the Soil and Potential Applications to Sorghum Breeding

Soil composition generally varies across a given unit of space (McNeil, 1980; Johnson and Eskridge, 2005; Corwin and Scudiero, 2016). Plant breeding programs typically attempt to account for this variability by using experimental designs that include replication and blocking of genotypes (Stroup et al., 1994; Hoshmand, 2006). This approach is generally acceptable at accounting for at least a portion of the variation that can be caused by spatial variability of the soil within a field; however, this technique is rarely assessed for its ability to account for soil variability per se (Stroup et al., 1994).

Measurement of apparent electrical conductivity (EC_a) of the soil is a simple, non-invasive and proximal method for mapping spatial variability of the soil (Johnson and Eskridge, 2005; Corwin and Scudiero, 2016). Multiple instruments are commercially available, but the EMI technique is truly non-invasive compared to EC_a measured using a DC current.

Soil EC_a values are affected by many factors that impact crop performance, but it is strongly correlated with water holding capacity (WHC), clay content, and soil moisture in non-saline soils (Rhoades et al., 1976; McNeil, 1980). Other studies have also shown a correlation between EC_a and crop yield (Robert et al., 1995a; b, Kitchen et al., 1999, 2003). Soil EC_a maps may be useful to plant breeders as an addition to statistical models or as an aid in a priori field placement of experimental trials (Johnson et al., 2005). As mentioned, previous research has investigated the relationship between soil EC_a and crop yields (Anderson-Cook et al.,; Johnson and Eskridge, 2005). In Johnson et al. (2005), an experiment was conducted in which an entire 32-hectare site was mapped using EC_a and was then partitioned into classes various. The classifications generated on EC_a were associated with the variability of soil properties including water content, clay content, and the presence of salts. Therefore, it is also possible that plant breeders may similarly be able to use this technology as a method by which they can identify and reduce the effect of spatial variability in their material or even as a replacement for replicated blocks in certain situations.

1.4 Estimation of Anthracnose Disease in Sorghum Using Unmanned Aerial Systems

Previous studies have shown the efficacy of remote sensing to estimate other biotic and abiotic stressors in crops, and these same concepts could apply to the phenotyping of crop diseases (Valasek et al., 2016; Caturegli et al., 2016; Yang et al., 2018). A common remote sensing technique to assess plant stress is normalized difference vegetation index, (NDVI), which is a rating that is often used as a measurement of plant health and photosynthetic integrity (Rouse et al., 1974; Jones, H. G., Vaughan, 2010). NDVI has been previously evaluated and implemented as an estimator of many parameters of interest including disease presence, photosynthetic activity, abiotic stress, and others (Rouse et al., 1974; Gamon et al., 1995; Anyamba and Tucker, 2005; Kumar et al., 2016).

Disease progress curves have been utilized by crop scientists for many years and have been applied in breeding programs previously (Néya and Le Normand, 1998; Hess et al., 2002; Li and TeBeest, 2009). The area under the disease progress curve, or AUDPC, can provide a quantitative assessment of disease severity within each individual genotype (Jeger and Viljanen-Rollinson, 2001). In much the same way that individual assessments of disease severity could be useful, disease progress curves generated using high-throughput techniques could produce new information for breeders.

Sorghum serves as an excellent case study by which to evaluate UAS for their ability to phenotype disease. While there are several diseases of significance in sorghum,

anthracnose (*Colletotrichum sublineola* P. Henn., Kabát and Bubák) is arguably the most important (Tebeest et al., 2004; Li and TeBeest, 2009; Moore et al., 2010). In sorghum, the disease infects all above ground plants of the part (stalk, leaves, peduncle, and panicle) producing visible damage (Warren, 1986; Moore et al., 2010). Anthracnose can cause sorghum yield reductions as high as 50% or more (Li and TeBeest, 2009). As such, anthracnose serves as an excellent case study to assess UASs for their ability to phenotype the same.

Genetic resistance is the primary means of controlling the disease. Because there are many different strains of *C. sublineola*, sorghum improvement scientists must continuously breed for resistance in their material (Leslie, 2002). Traditional methods of phenotyping sorghum for the presence and severity of the disease in the field are ultimately effective but are also quite laborious and are prone to a significant amount of measurer error due to their inherent subjectivity of the rating. Thus, the objectives of this study were to i) evaluate the relationship between remotely-sensed NDVI data taken using a fixed-wing UAS and ground-truth anthracnose disease rating, and to ii) evaluate each of the various measurements in this study for their relationship to final grain yield.

2. ESTIMATION OF BIOMASS YIELD AND PLANT HEIGHT IN SORGHUM USING UNMANNED AERIAL SYSTEMS

2.1 Synopsis

Sorghum (*Sorghum bicolor* L. Moench) has been traditionally been grown as a forage and it has recently been bred specifically for conversion to a renewable fuel. Producers of forage and bioenergy sorghums both value high biomass yield. Traditional breeding approaches require significant investments in capital and labor. Unmanned aerial systems (UAS) have demonstrated utility in estimating several agronomic traits. Herein we present a statistical analysis and evaluation of UAS for their ability to estimate biomass yields and plant heights in forage and bioenergy sorghum hybrids. We find that the correlation between ground-truth estimates of biomass yield and one of the UAS estimates of the trait (canopy volume, VOL) ranged from $r = 0.69 – 0.85$, dependent upon the environment. In addition, the UAS estimates of plant height ($r = 0.74 – 0.99$) and biomass yield ($R = 0.91 – 0.93$) were as repeatable as the corresponding ground-truth methods. Finally, multi-temporal growth curves generated herein have increased potential utility to breeders, because they include two different traits that both serve as components of biomass yield. There appears to be potential in using UAV estimates in forage and bioenergy sorghum breeding programs.

2.2 Introduction

Breeding crops for biomass productivity has been important even before the adoption of scientific approaches to breeding began to develop over a century ago. Today, breeding for biomass productivity takes many forms; in sugarcane, sugar and bioenergy production are priorities, and in alfalfa, yield, quality, and stand persistence are important. Regardless of the biomass crop, biomass yield is a common and critically important trait. Plant breeders are constantly working towards improving biomass yield so that they are a more attractive forage, lumber or alternative for energy producers. However, alongside other notable factors, phenotyping has heretofore remained one of the primary bottlenecks in modern breeding programs (Furbank and Tester, 2011). Therefore, high-throughput methodologies must be developed to ensure that bioenergy crops can be efficiently improved (Walter et al., 2012; Araus and Cairns, 2014).

As mentioned, there are several different crops that are grown for their superior biomass properties. Perhaps one of the most promising of these is sorghum (*Sorghum bicolor* L. Moench) (Rooney et al., 2007; Calviño, 2012). At non-Equatorial latitudes, photoperiod-sensitive sorghum maintains vegetative growth for a much longer time than photoperiod-insensitive genotypes resulting in higher biomass production (Rooney and Aydin, 1999; Murphy et al., 2011). Of the many traits of interest, perhaps the most important are plant height and biomass yield, wherein the former is often an enormously important component of the latter (Machado et al., 2002; Salas Fernandez et al., 2009; Yin et al., 2011; Roth and Streit, 2018). Thus, breeding for the improvement of both forage and bioenergy sorghum is quite different from the approach that would be

followed when breeding for increased grain yields in the same crop. Additionally, the traits mentioned can be quite laborious to phenotype, thereby creating the impetus for the evaluation of high-throughput techniques (Furbank and Tester, 2011; White et al., 2012).

One potential method that breeders can use to increase the efficiency and accuracy of their efforts to improve their crops is to use remote sensing via unmanned aerial systems (UAS) and photogrammetry (Colomina and Molina, 2014; Shi et al., 2016; Roth and Streit, 2018; Shafian et al., 2018). Several recent studies have already demonstrated that photogrammetry can potentially be used to predict biomass or height in several crops including maize (*Zea mays*), wheat (*Triticum aestivum*), sorghum, sugarcane (*Saccharum officinarum*), barley (*Hordeum vulgare*), soybean (*Glycine max*), poppy (*Papaver somniferum*), and grapes (*Vitis vinifera*) (Grenzdörffer,; Bendig et al., 2015; Watanabe et al., 2017; De Souza et al., 2017; Matese et al., 2017; Iqbal et al., 2017; Yue et al., 2017; Pugh et al., 2017; Malambo et al., 2018; Roth and Streit, 2018). Sorghum scientists could use cameras mounted on UAS to predict total biomass yield in sorghum, and temporal monitoring of the total growth and development of canopy height and volume (Cooper et al., 2016; Pugh et al., 2017).

Researchers that use UAV methodology could also elucidate details that were previously difficult or impossible to describe using traditional techniques. For example, constructing high-resolution multi-temporal growth curves is a novel use for UAS technology, wherein plant breeders could investigate the overall vigor of a genotype over the entire growing period (Cooper et al., 2016; Pugh et al., 2017). Additional

applications could include delineating maturity subgroups within breeding populations, determining optimum timepoints for which to harvest each genotype, and assessing overall response to stresses (Cooper et al., 2016; Pugh et al., 2018).

Given this background, the objectives of this study were i) to evaluate UASs for their ability to predict biomass yield and plant height in bioenergy sorghum, ii) to compare the ability for ground-truth measurements and UASs to determine the amount of genotypic variation in a population, and iii) to construct multitemporal growth curves using UAS data to tease out maturity differences as well as variation in growth patterns between various genotypes.

2.3 Materials and Methods

2.3.1 Germplasm and Experimental Design

A test composed of 15 sorghum genotypes consisted of biomass material including six commercial forage hybrids, six open pedigree biomass and forage hybrids developed by Texas A&M University, two sweet sorghum cultivars, and one bioenergy sorghum cultivar. These bioenergy and forage genotypes were expected to show variation for several traits of interest, including plant height and biomass yield which are two key traits for biomass sorghum producers (Table 1). Additionally, the genotypes used in this study were expected to show variation for their maturity date, or flowering date (Table 1).

The biomass soghum test was planted in a randomized complete block design (RCBD) with four replications and included two growing seasons (Hoshmand, 2006). In 2016, this test was planted with two-row plots in Corpus Christi, TX (CC16) on March

29 and College Station, TX (CS16) on March 26. In 2017, the test was planted with four-row plots only in College Station, TX (CS17) on April 1. Each individual research plot measured 6.7-m long and included a 1.22-m alley in both locations. Standard agronomic practices for bioenergy sorghum were applied.

2.3.2 Field Phenotyping for Ground-truth Validation of Plant Height, Bioenergy Quality Traits, and Biomass Yield

Plant height (cm) was measured using different methods based on the growth stage of the plants in each plot (Pugh et al., 2018). Plants with panicles which had not yet emerged from the whorl were measured from the bottom of the plant to the highest point, or apex, of the plant. The apex was defined the highest point, including leaves, that could be accurately measured by researchers on the ground and usually was congruous to measuring the height of the canopy (Pugh et al., 2018). Plants that had emerged from the whorl were measured from the ground to the panicle tip. Though it used a different morphological attribute on the plants, this measurement was nonetheless congruous to the apex measurement used on plants with panicles that had not yet emerged. Measurements of height represented a mean estimate across an entire research plot. The maximum height of each plot, whether it was at the apex or the panicle tip, was used as the "ground-truth" value (Pugh et al., 2018). Biomass yield (kg ha^{-1}), or BMYLD, was estimated by harvesting every plant within the two and four-row plots, respectively, using a plot harvester at the end of the growth period.

2.3.3 Flight Planning, Image Acquisition, and Image Analysis

Flights were conducted by the same respective teams in each location, and in the same manner, as described in Pugh et al. (2018). In addition, maximum percentile (MAX) and 95th percentile (P95) plant heights were estimated using the same methodologies described in that study (Pugh et al., 2018). The same RGB imagery that was obtained and used to generated plant height estimates was subsequently used to calculated canopy volume (VOL) and canopy cover (COV) estimates.

Prior to calculating the volume and canopy cover for CS16 and CS17, each generated point cloud dataset was normalized to above-ground level (AGL) by subtracting the corresponding ground elevation from a generated digital elevation model (DEM). For each flight date, the resulting AGL point cloud was then split into sub-point clouds covering the extents of each biomass sorghum plot. Using each of these sub-point clouds, VOL and COV were estimated for each plot.

Canopy volume is a product of the canopy surface area that was covered and the canopy height, herein presented as the plant height. However, to estimate VOL, the differences in plant height and COV were accounted using a grid-based approach. Each sub-point cloud was thinned to capture the top of canopy points by retaining only the highest point in a 7.5 cm by 7.5 cm cell. The point density for all the datasets was generally greater than 2000 points per square meter. Thus, a 7.5 cm grid presented a good compromise between reducing the amount of data for the following computation and maintaining the 3D structure of the canopy. The total volume over the plot was then determined by summing contributions from all the cells. Non-foliage points were

excluded based on excess green index (ExG) values using the binary Otsu thresholding process. For the calculation of COV, each sub-point cloud was thinned as described above. The canopy cover was then estimated as the percentage of the number of foliage points with respect to the total number of points over the entire plot.

For the CC16 location, binary classification (canopy vs. non-canopy area) of the RGB orthomosaic images was completed by applying the Canopeo algorithm (Patrignani and Ochsner, 2015). Canopy cover was calculated by dividing the canopy area by the total plot area. For this process, a rectangle polygon was adopted which had a uniform size for each row in the field. For VOL, canopy height models (CHM) were generated by subtracting the digital terrain model (DTM) from the digital surface model (DSM) for each UAS flight (Dandois et al., 2010). Canopy volume was calculated from the CHM by multiplying the canopy height with the pixel area within each individual plot.

2.3.4 Data Analysis and Statistics

Sorghum ground-truth measurements and UAS estimates were checked for outliers in JMP Pro 12.2.0 software (SAS Institute Inc., 1989 - 2017). Correlations were conducted between ground-truth measurements and their associated UAS estimates in JMP Pro 12.2.0 and were reported as Pearson's correlations coefficients (*r*). Restricted maximum likelihood analysis (REML) was conducted within environments using Fit Model (all random) in JMP. The statistical model that was used for this analysis was

$$Y = \alpha_i + \beta_j + \gamma_t + \delta_k + \varepsilon$$

, where Y = biomass yield or plant height, α = genotype (i), β = replication (j), and ε = error (D'Agostino et al., 2006). Effects with negative variance components were

removed from the model. The percentage of the total variation explained by genotype was calculated using this model as was repeatability. Repeatability (*R*) estimates were calculated using the equation:

$$R = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}}$$

, where *R* = the repeatability score, σ_g^2 = the genotypic variance, σ_e^2 = the error variance, and *r* = the number of replications (Nakagawa and Schielzeth, 2010). Repeatability is a metric that uses the variance of each component to give researchers an indication about the general consistency of their techniques. It is calculated similarly to but is also distinct from heritability (H^2) since the bioenergy sorghum population used in this study did not have a familial structure. The percentage of variation explained as well as repeatability estimates were calculated for each phenotype within each flight date to determine the consistency and power of the technology across the growing season.

Pearson's correlation coefficients (*r*) were calculated using JMP Pro 12.2.0 software. Correlations were conducted for the Bioenergy test in both years and locations. Least squares means (LSMeans) were calculated using the same model noted above, except that genotype and replication were considered fixed effects. The LSMeans were used to construct high-resolution multitemporal growth curves for UAS-estimated canopy volume. To create a smooth growth curve, the Smoother function was used in JMP 12.2.0 software with a lambda value of 0.05.

2.4 Results and Discussion

2.4.1 Correlations Between Ground-truth and UAS-based Measurements of Biomass Yield and Plant Height

Overall, the correlations between VOL and BMYLD were strong, particularly in CC 16 and CS 17 (Table 1). However, the correlation was lower in CS 16 (r = 0.69) (Table 1). The specific reason for this is not known, but it could be due to some of the same factors previously described in other studies (Pugh et al., 2017; Malambo et al., 2018). It is also possible that the UAS methodology used in CS performed better when being used with four-row plots instead of two-row plots. Regardless, the correlation between VOL and BMYLD was high in all three locations (Table 1). In stark contrast, COV was not nearly as strongly correlated with BMYLD (Table 1). The correlations ranged from low-moderate in CS 16 and CC 16 to very low (r = 0.18) in CS 17. These results strongly indicate that VOL is an excellent predictor of BMYLD in sorghum and should be utilized by breeders in their programs an should certainly be used in lieu of COV (Table 1).

Table 1. Correlations Between UAS Estimates and Biomass Yield in Sorghum Pearson's correlation coefficients (*r*) between canopy volume (VOL), canopy cover (COV), and biomass yield (BMYLD) in bioenergy sorghum. These correlations were conducted in three environments: Corpus Christi in 2016 (CC 16), College Station in 2016 (CS 16), and College Station in 2017 (CS 17).

	CC 16	**CS 16**	**CS 17**
BMYLD vs. VOL	0.85***	0.69***	0.81***
BMYLD vs. COV	0.45**	0.41**	0.18
VOL vs. COV	0.44**	0.66***	0.54***

** Significant at $p < 0.01$
*** Significant at $p < 0.001$

The relationship between GT measurements of plant height and those derived from UAS (MAX, P95) has been demonstrated in grain sorghum or in smaller field plots previously (Watanabe et al., 2017; Pugh et al., 2017; Malambo et al., 2018). Interestingly, the correlations between GT estimates and UAS estimates were like those described for the Early Generation material in Pugh et al. (2018) but were much more heavily contrasted and did not exhibit any ambiguity towards the end of the growth period. Thus, P95 was definitively a superior estimate to use for plant height in this study (Table 2). In fact, P95 was superior in all three environments, and in every measurement date within those three environments (Table 2). The correlations were the highest in CS 16, reaching a Pearson's value of almost 1.00 in one case (69 days after planting, or DAP). These results show that P95 is an obvious best choice for estimating height in bioenergy material (Table 2).

Table 2. Correlations Between UAS Estimates and Plant Height in Sorghum Pearson's correlation coefficients (*r*) between the maximum and 95th percentiles of UAS point cloud data (MAX and P95, respectively) and ground-truth (GT) estimates of plant height. These correlations were conducted in three environments: Corpus Christi in 2016 (CC 16), College Station in 2016 (CS 16), and College Station in 2017 (CS 17). Measurements were performed multiple times, and measurement dates are shown as the number of days after planting (DAP) that the measurements were recorded.

CC 16		
DAP	**GT vs. MAX**	**GT vs. P95**
29	0.42**	0.43**
48	0.65***	0.66***
59	0.94***	0.96***
70	0.77***	0.77***
86	0.84***	0.86***

CS 16		
DAP	**GT vs. MAX**	**GT vs. P95**
40	0.61***	0.64***
69	0.88***	0.97***
73	0.77***	0.93***
82	0.80***	0.91***
90	0.72***	0.84***
97	0.69***	0.76***
104	0.58***	0.71***

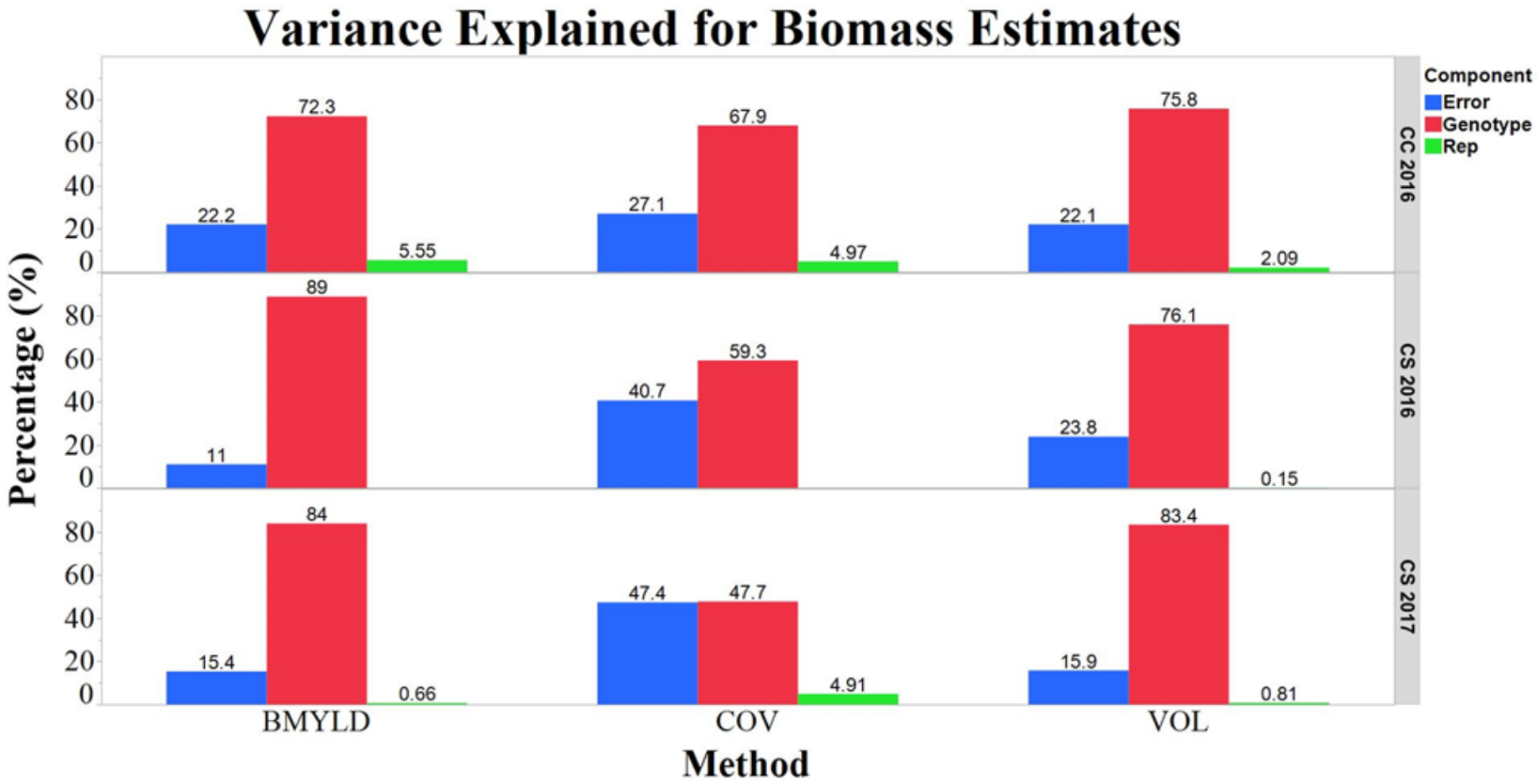
CS 17		
DAP	**GT vs. MAX**	**GT vs. P95**
54	0.69***	0.71***
59	0.82***	0.89***
76	0.69***	0.82***
97	0.69***	0.74***
104	0.64***	0.72***

** Significant at $p < 0.01$
*** Significant at $p < 0.001$

2.4.2 Variance Explained and Repeatability Estimates for Biomass Yield and Plant Height

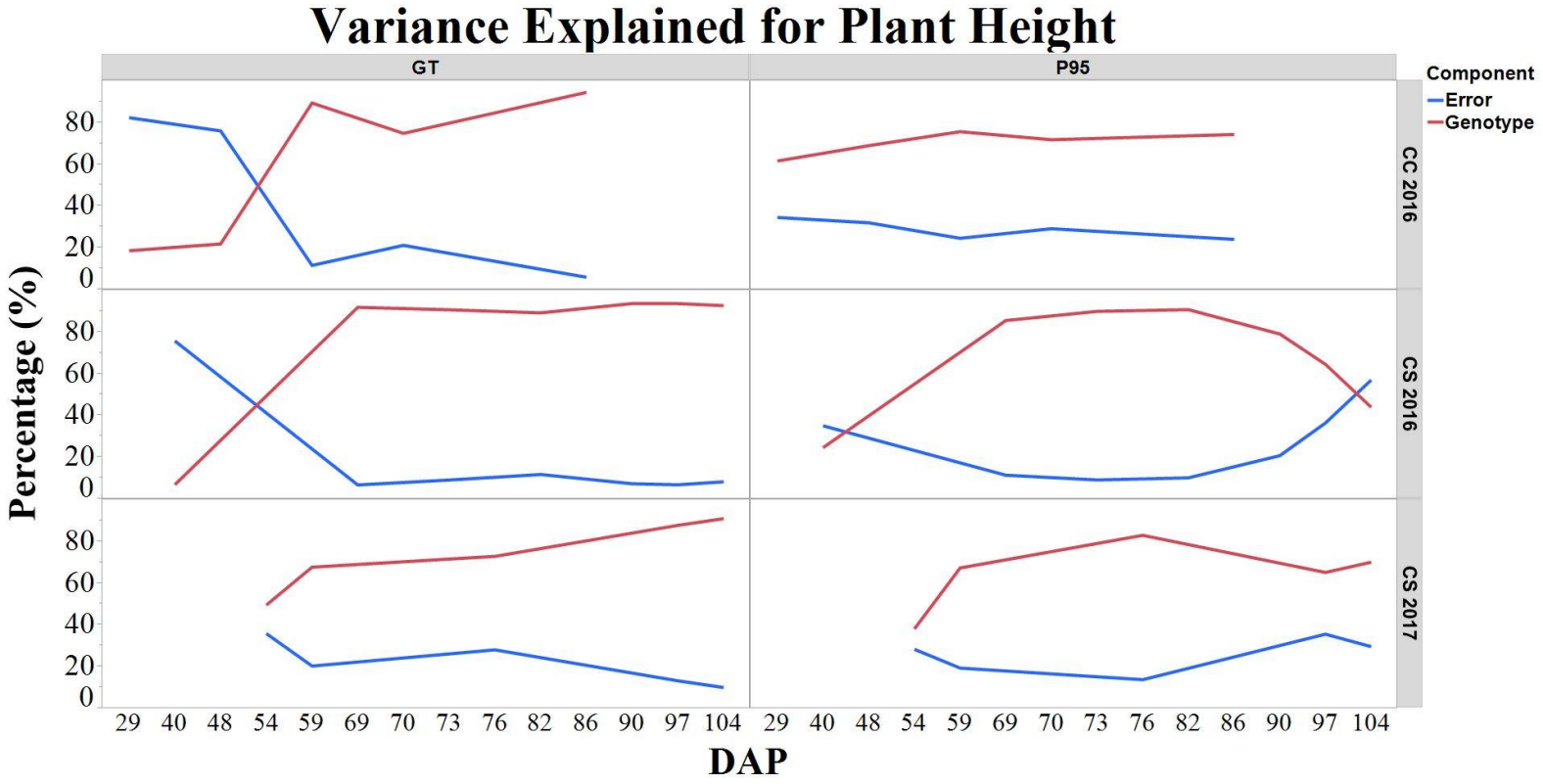
The results showed that VOL can tease out just as much genotypic variance as the traditional methods, with the notable exception of CS 16 as described for the correlations above (Figure 1). Regardless, the other two environments of CC 16 and CS 17 showed the same results, except that it did not matter which of the two methods was used (Figure 1). Again, in contrast, COV was an inferior method both for teasing out genotypic variation as well as for minimizing error (Figure 1). In fact, the error variance rose quite significantly when using that method, and the genotype variance plummeted. These results show that COV is not only very weakly correlated with BMYLD, it is also not able to differentiate between genotypes accurately; thus, COV is still not likely to be a reliable tool for plant breeders interested in bioenergy sorghum improvement (Figure 1).

Figure 1. Percentage of Variance Explained for UAS and Ground-truth Estimates of Biomass Yield. The percentage of variation that was explained for each environment and for each method of estimating biomass yield in sorghum. Methods analyzed included ground-truth estimates of biomass yield (BMYLD), as well as UAS-derived canopy volume (VOL) and canopy cover (COV). The statistical model used for this analysis was $Y = \alpha_i + \beta_j + \gamma_t + \delta_k + \varepsilon$, where α = genotype (i), β = replication (j), and ε = error. This analysis was conducted in three locations in Texas: Corpus Christi in 2016 (CC 16), College Station in 2016 (CS 16), and College Station in 2017 (CS 17).



For plant height, the ground-truth (GT) method of measuring plant height with a height stick was compared to the 95th percentile of the UAS height data. Since MAX was shown to be markedly inferior in all the correlations, it was excluded from this comparison. Overall, the variation explained was very similar between the methods in all three of the locations, although P95 was much more reliable early on in CC 16 (Figure 2). These results were not especially different from those demonstrated previously in grain sorghum trials (Malambo et al., 2018; Pugh et al., 2018). However, it should be noted that P95 did not reach final levels of variance explained for genotype that were quite as high as when using GT, especially near the end of the growth period (Figure 2). Again, this was not entirely unexpected since the same trends were also observed in our previous study (Pugh et al., 2018). Nonetheless, the trends that were observed in this study strongly suggest that P95 is an acceptable and efficient method to tease out genotypic variation for plant height in bioenergy breeding trials.

Figure 2. Percentage of Variance Explained for UAS and Ground-truth Estimates of Plant Height. The percentage of variation that was explained for each environment for ground-truth and UAS estimates of plant height in sorghum. Methods analyzed included ground-truth estimates of plant height (GT) as well as the 95th percentile of UAS-derived point cloud data (P95). The statistical model used for this analysis was $Y = \alpha_i + \beta_j + \gamma_t + \delta_k + \varepsilon$, where α = genotype (i), β = replication (j), and ε = error. Percentages of Genotype and Error are plotted over the course of the growing period, shown as the progression of the number of days after planting (DAP). This analysis was conducted in three locations in Texas: Corpus Christi in 2016 (CC 16), College Station in 2016 (CS 16), and College Station in 2017 (CS 17).



The previous results show that a sorghum breeder could likely make determinations within their bioenergy material strictly using the VOL and P95 metrics (Figure 1, Figure 2). In similar fashion, the repeatability estimates for VOL and BMYLD were very similar across all three environments (R = 0.93 – 0.97), while COV was considerably lower (Table 3). The similarity of the repeatability estimates was expected, as repeatability estimates are based entirely upon the relative amounts of genotypic and error variance. Nonetheless, those estimates corroborate the same trend seen previously: that COV is not a very effective way to estimate biomass yield in bioenergy sorghum. The repeatability scores for the three different plant height measurement techniques were all high throughout most of the growing period, although the estimates were lower for GT early in the growth season and, conversely, were lower for the P95 and MAX metrics at the end of the growth period (Table 4). These results are also like those that were described in Pugh et al. (2018) when using the same measurement techniques on grain sorghum, although the decreased values at the beginning and end of the growing season were more pronounced in this study (Table 4).

Table 3. Repeatability Estimates for UAS and Ground-truth Estimates of Biomass Yield.
Repeatability (*R*) estimates are included for ground-truth estimates of biomass yield (BMYLD), as well as UAS-derived estimates of canopy volume (VOL) and canopy cover (COV). Repeatability estimates were calculated using the equation $R = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}}$, where *R* = the repeatability score, σ_g^2 = the genotypic variance, and σ_e^2 = the error variance. These repeatability estimates were calculated for three locations in Texas: Corpus Christi in 2016 (CC 16), College Station in 2016 (CS 16), and College Station in 2017 (CS 17).

	CC 16	**CS 16**	**CS 17**
BMYLD	0.93	0.97	0.96
VOL	0.93	0.93	0.95
COV	0.91	0.87	0.80

Table 4. Repeatability Estimates for UAS and Ground-truth Estimates of Plant Height. Repeatability (*R*) estimates are included for ground-truth (GT) as well as the maximum and 95th percentiles of UAS-derived point cloud data (MAX and P95, respectively). Repeatability estimates were calculated using the equation $R = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}}$, where *R* = the repeatability score, σ_g^2 = the genotypic variance, and σ_e^2 = the error variance. Estimates are shown for each flight date during the growth period, shown as the progression of the number of days after planting (DAP) that the measurements were taken. These repeatability estimates were calculated for three locations in Texas: Corpus Christi in 2016 (CC 16), College Station in 2016 (CS 16), and College Station in 2017 (CS 17).

CC 16			
DAP	**GT**	**MAX**	**P95**
29	0.47	0.88	0.88
48	0.53	0.88	0.89
59	097	0.91	0.93
70	0.93	0.89	0.91
86	0.99	0.91	0.93

CS 16			
DAP	**GT**	**MAX**	**P95**
40	0.25	0.82	0.74
69	0.98	0.95	0.97
73	0.98	0.84	0.98
82	0.97	0.90	0.97
90	0.98	0.81	0.94
97	0.98	0.79	0.88
104	0.98	0.59	0.75

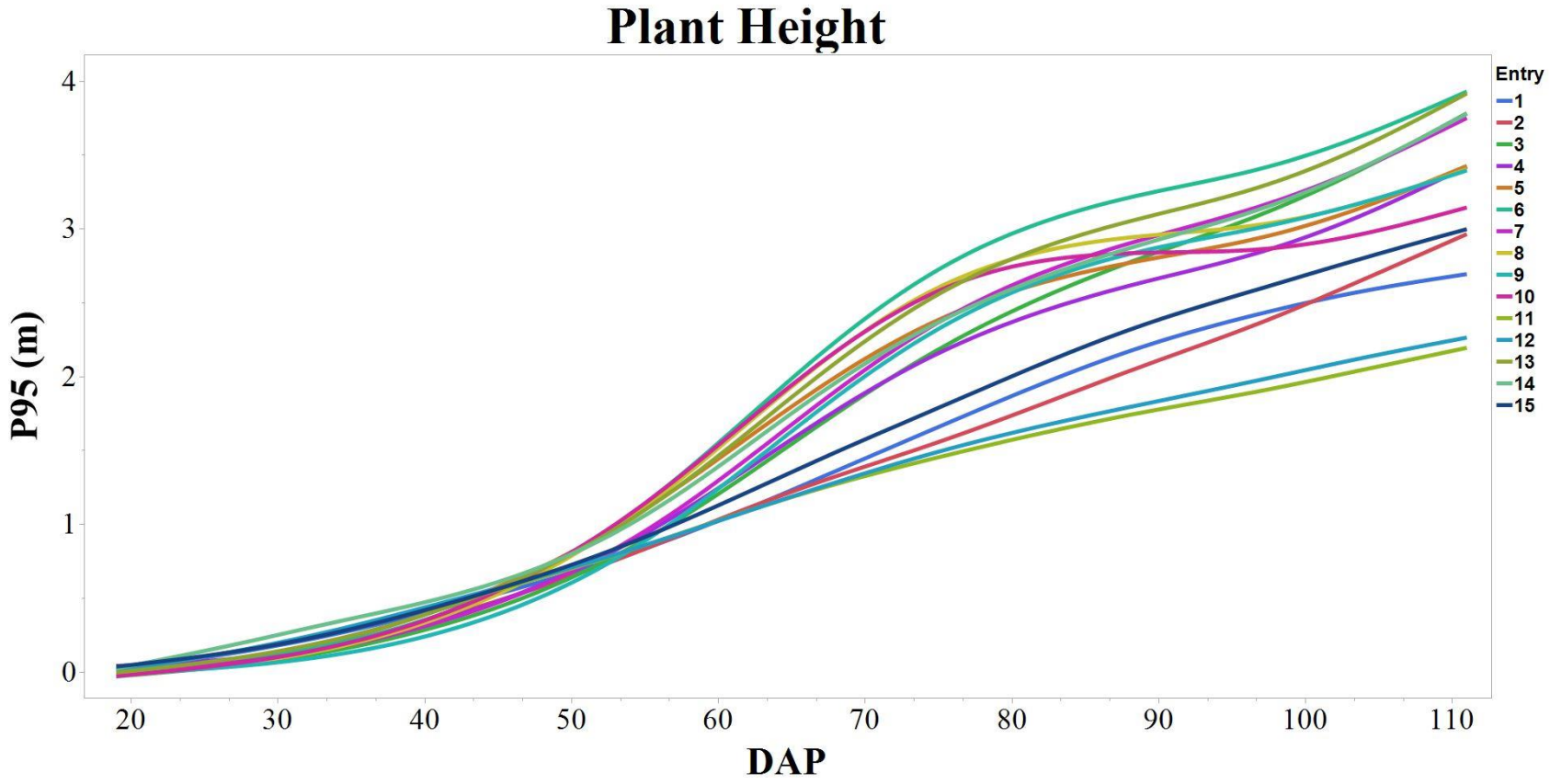
CS 17			
DAP	**GT**	**MAX**	**P95**
54	0.85	0.75	0.84
59	0.93	0.83	0.93
76	0.91	0.83	0.96
97	0.96	0.86	0.88
104	0.97	0.78	0.91

2.4.3 Multitemporal Growth Curves for Biomass Yield and Plant Height

One phenotype that can potentially be described more efficiently and effectively using UAS is that of the temporal growth curve. This technique can be used to better define growth patterns of bioenergy sorghum as well as to tease out other phenotypic characteristics that can cause changes in a genotype's growth pattern (Cooper et al., 2016; Pugh et al., 2018). The 95th percentile of plant height (P95) measurement was plotted on a high-resolution multitemporal growth curve using UAS data from the CS 17 trial (Figure 3). When using this methodology, the fifteen different genotypes used in this study exhibited distinct growth patterns from each other (Figure 3). These differences in growth patterns allowed us to tease out interesting phenomena that would be difficult to identify without using a growth curve, such as various "cross-over" events where one genotype would overtake another in overall P95 values, as well as the characterization of photoperiod sensitivity. One example of such an occurrence was at about 100 DAP, when Genotype 7 was displaced by Genotype 4 and ended up having a lower final P95 value (Figure 3). Genotype 7 had a higher estimate than Genotype 4 for the rest of the growing season (Figure 3). In fact, Genotype 7 exhibited a small decrease near the end of the season. The reason for this is not known, although it has been speculated upon previously (Pugh et al., 2018). There were several other examples of such cross-over events during the growth-period, and the multitemporal growth curve helps to illustrate these changes in ranking as the season progresses (Figure 3). Cross-over information may be valuable not only to help plant breeders determine which genotypes will have superior yield or respond to environmental stressors, but it may also

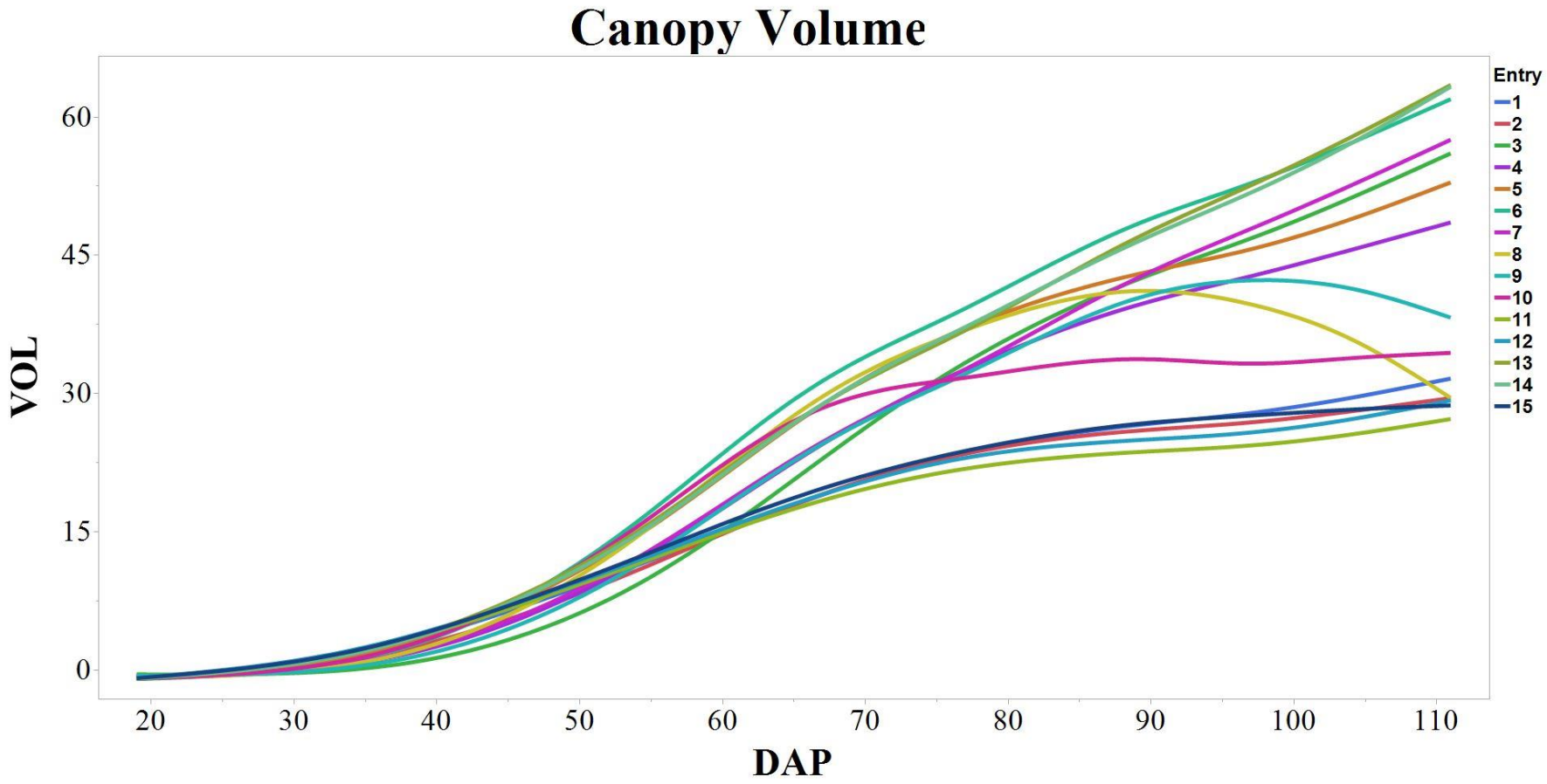
elucidate what the optimal harvest date is for a given genotype by observing when a growth curve reaches its zenith (Figure 3).

Figure 3. Temporal Growth Curve for Plant Height in Biomass Sorghum. Representation of a UAS-derived growth curve for plant height (m) fit during the period of growth for a bioenergy sorghum trial in College Station in 2017, shown as the progression of days after planting (DAP) of each measurement). Each color on the curve denotes the 95th percentile of the structure-from-motion (SfM) data for a different genotype, with fifteen genotypes total.



For canopy volume (VOL), the growth curves were very similar to those for plant height, which was expected since the two traits are closely related (Figure 4). Thus, in a similar fashion to P95, there were multiple temporal shifts that occurred during the growth period. However, there were some notable differences. For example, Genotypes 8 and 9 saw a large decrease in VOL values near the end of the growth period, but this same decrease was not observed in their growth curve for plant height (Figure 4). It is not known why this occurred, although it could potentially be attributed to some of the same explanations that were also proposed for plant height in Pugh et al. (2018). This reduction in VOL near the end of the growth period was only observed in those two genotypes, although marked differences in their rate of increase by that point of growth were also observed among the other genotypes in the study (Figure 4). It is possible that these differences could have been influenced by other ancillary phenotypic characteristics of importance to bioenergy sorghum breeders; however, this explanation will remain speculative until further studies can better characterize the interactions between various phenotypic traits within the material.

Figure 4. Temporal Growth Curve for Canopy Volume in Biomass Sorghum. Representation of a UAS-derived growth curve for canopy volume (VOL) fit during the period of growth for a bioenergy sorghum trial in College Station in 2017, shown as the progression of days after planting (DAP) of each measurement). Each color on the curve denotes the VOL estimate for a different genotype, with fifteen genotypes total.



One of the key considerations for biomass sorghum production is whether it is photoperiod sensitive (PS) or photoperiod insensitive (PI) (Rooney et al., 2007). Photoperiod sensitive material is preferable for biomass producers in the United States since it will have drastically delayed flowering and, thus, much longer periods of vegetative growth (Major et al., 1990; Rooney and Aydin, 1999; Rooney et al., 2007). Delayed maturity is advantageous to biomass producers, since greater amounts of vegetative growth will result in increased biomass yields and more fuel (Rooney et al., 2007). Conversely, it is more desirable for other ideotypes, such as grain sorghum, to be PI. Using multitemporal growth curves, it is possible to observe the results of photoperiod sensitivity, or lack thereof, in bioenergy sorghum material (Figure 4). For example, Genotype 15, which is a PI sorghum line, reached a lower maximum height than the rest of the material (Figure 3, Figure 4). Conversely, Genotype 7 (Graze-N-Bale) is a PS sorghum line that did not flower during its growth period, and its growth curve continued to increase throughout the season (Gill et al., 2014) (Figure 3, Figure 4). Other examples of this trend can be observed as there are several PI and PS representatives in the population. Genotypes 1, 2, 8, 9, and 10 were all some degree of PI, while the rest of the genotypes were PS. Further research will be required to elucidate the precise reasons for the differences between growth curves for canopy volume and plant height in a given set of bioenergy material, as well as their relationship to maturity. Nevertheless, the utility of this technique is as clear for this metric as it is for plant height, and it will allow plant breeders to make more informed decisions about which material to advance within their respective programs.

2.5 Conclusions

In conclusion, the results presented herein clearly indicate that UAS-mounted sensors can be effective estimators of biomass yield and plant height in bioenergy and forage sorghum. The P95 and VOL measurements are excellent predictors of plant height and biomass yield, respectively. In addition, the UAS-based measurements can tease out similar levels of genotypic variation while minimizing error, as was also shown in previous studies that investigated UAS for their ability to predict these phenotypic traits in grain sorghum or performed their experiments on smaller plots (Watanabe, 2017; Malambo, 2018; Pugh et all, 2018). Finally, the potential utility of using high resolution multitemporal growth curves has been shown in the current study to be just as much of a potential asset as it was in the prior study involving grain sorghum trials, albeit for additional reasons (Pugh et al., 2018). However, it is important that researchers take care when choosing their approach to measuring these various traits, since the MAX height percentile and the COV estimates were not nearly as useful for predicting their respective final phenotypic values in this material. Nevertheless, the utility of using a rotary-winged UAS for predicting biomass yield and plant height in bioenergy sorghum should not be understated. Further studies will be required to truly assess the impacts of implementing these technologies on a large scale, but the results herein show that UAS are a viable alternative to traditional phenotyping methodologies in sorghum breeding programs.

3. A STATISTICAL EVALUATION OF REPLICATED BLOCK DESIGNS AND SOIL ELECTRICAL CONDUCTIVITY GRIDS IN SORGHUM PERFORMANCE TRIALS

3.1 Synopsis

Success of variety selection trials have always been susceptible to the efficiency at which the experimental design can remove any effect of spatial autocorrelation associated with environmental factors. Blocking in a randomized design is one way of accounting for this spatial variability. Another way is to have a model of the environmental variability. Measures of soil variability could be useful to represent spatial structure in a trial, if the soil is the main factor creating spatial variability in trail results. Soil apparent electrical conductivity (EC_a) measurements can be collected rapidly and non-invasively, and have been well documented to be able to map soil variability at the meter scale. We present a statistical evaluation that compares the effectiveness of the traditional replicated block designs, with spatially explicit soil EC_a measurements. Soil EC_a, grain yield and plant height were measured for six sorghum hybrid evaluation trials across Texas in 2017. Three linear models were tested to determine the presence or absence of spatial autocorrelation of model residuals within each performance trial. Moran's I tests on model residuals showed that neither method was consistent effective in accounting for spatial variability. Blocking was more effective at one site for both plant height and grain yield, while EC_a data were more effective at two sites for grain yield only. Based on these results, and the relatively low cost of using both methods together, we propose that plant breeders interested in

addressing spatial autocorrelation in models from trials results may consider using both methods and select the best model post hoc.

3.2 Introduction

Projected population growth as well as a global climate change have created the impetus for more efficient techniques to be implemented in crop improvement programs (Godfray et al., 2010; Ray et al., 2012, 2013; Cairns et al., 2013; Araus and Cairns, 2014). Sorghum (*Sorghum bicolor* L Moench) stands as an important crop to meet these growing food, bioenergy, and forage demands. To that end, sorghum researchers are continually validating and implementing new ways to improve the efficiency of their programs. Non-invasive techniques, including the proximal and remote sensing of soil properties, stand as attractive potential opportunities to rapidly characterize the spatial variability of a growing environment. A plant breeder may use these technologies to predict performance of various genotypes in their trials (Furbank and Tester, 2011; Araus and Cairns, 2014; Pugh et al., 2018).

Soil composition generally varies across a given unit of space (McNeil, 1980; Johnson and Eskridge, 2005; Corwin and Scudiero, 2016). Plant breeding programs typically attempt to account for this variability by using experimental designs that include replication and blocking of genotypes (Stroup et al., 1994; Hoshmand, 2006). This statistical approach has need accepted by most plant breeders for accounting for and adequate portion of the variation that may be caused by soil variability; however, blocking is rarely assessed for its ability to account for soil variability per se (Stroup et al., 1994). Apparent soil electrical conductivity (EC_a) is a simple, non-invasive and

proximal method for mapping spatial variability of the soil (Johnson et al., 2005; Corwin and Scudiero, 2016). Multiple instruments are commercially available, but the electromagnetic induction (EMI) technique is truly non-invasive compared to the DC electrical conductivity method using coulters drug through the soil surface. Both techniques allow for a rapid survey of soil variability (Johnson et al., 2005; Corwin and Scudiero, 2016).

Soil EC_a values are affected by many factors that impact crop performance, but it is strongly correlated with water holding capacity and clay content, especially in well drained-non-saline agricultural soils (Rhoades et al., 1976; McNeil, 1980). In high-clay Texas Vertisols, soil ECa tends to be most correlated with calcium carbonate concentration in eth soils (Neely et al., 2016). Agronomic studies have also shown a correlation between EC_a and crop yield (Robert et al., 1995a; b, Kitchen et al., 1999, 2003); however the sign and magnitude of correlation depends on annual weather variables. For example, a high clay location in eth field can be strongly and positively correlate din a dry growing season but negatively correlated to yield in a year when the spring is wet and seeding growth is inhibited by the soil being too wet.

Because of the association between ECa and soil properties that also drive a plants yield performance, soil EC_a maps may be useful to plant breeders as an addition to statistical models or as an aid in a priori field placement of experimental trials (Johnson et al., 2005). As mentioned, previous research has investigated the relationship between soil EC_a and crop yields (Anderson-Cook et al., 2002; Johnson and Eskridge, 2005). In Johnson et al. (2005), an experiment was conducted in which an entire 32-hectare site

was mapped using EC_a and was then partitioned into classes various. The classifications generated on EC_a were associated with the variability of soil properties including water content, clay content, and the presence of salts. Therefore, it is also possible that plant breeders may similarly be able to use this technology as a method by which they can identify and account for effect of spatial patterns in their field trials or even as a replacement for replicated blocks in certain situations.

To assess the effectiveness of using interpolated EC_a grids as a tool for plant breeders, data were collected and analyzed within six replicated sorghum performance trials in Texas in 2017. The objectives of this study were i) to quantify the level of spatial autocorrelation present in yield and plant height data at each trial location, and ii) to assess the effectiveness of two linear models, one using traditional blocking via replication and the other using EC_a grids, that can be used to account for spatial autocorrelation.

3.3 Materials and Methods

3.3.1 Experimental Design and Germplasm

The test subjects of this study were grain sorghum evaluation trials which were composed of different groups of advanced hybrid genotypes. The specific purpose of these performance trials is to evaluate sorghum hybrids for their relative grain yield and agronomic productivity in different production environments. These trials were in Texas and included Danevang, Greenville, Gregory, Mexia, Perryton, and Plainview. Environments contained either 30 (Danevang, Gregory, Greenville) or 40 genotypes (hybrids) (Plainview, Mexia, Perryton) arranged in two-row plots replicated four times

in a randomized complete block design (RCBD) (Hoshmand, 2006). The width of the rows varied by location, but were either 0.76 or 1.02 m, and the plot length was 9.14 m (after a 1.22 m alley was cut). Most of the trials were unirrigated, except for Perryton and Plainview which were irrigated (Table 5). Since variation in the soil was a primary interest in this study, the Official Soil Series Descriptions (OSDs) were also identified for each performance trial (Table 5) (Soil Survey Staff, USDA-NRCS).The dominant soil type in each location was Lake Charles Clay in Danevang, Houston Black Clay in Greenville, Raymondville Clay Loam in Gregory, Ferris-Heiden Complex in Mexia, Sherm Clay Loam in Perryton, and Pullman Clay Loam in Plainview (Table 5). Standard agricultural practices for sorghum grain production were practiced at each of these trials.

The plant height data for this experiment was manually collected at the end of growth and development using a measuring stick. All the locations in this study were harvested and the total grain yield was concurrently estimated using a John Deere 3300 plot combine equipped with the HarvestMaster Grain Gauge. Field and harvest notes were then compiled by the Texas Crop Testing group and were released following their respective analyses (Schnell et al., 2017).

3.3.2 Collection of EC_a Data and Interpolation of EC_a Grids

Soil EC_a $(mS\ m)^{-1}$ was measured using an EM38-MK2 (Geonics Limited, Mississauga, Ontario, Canada ®), which was manually walked throughout the entire performance test in each location. The EM38 was used in the vertical orientation to ensure maximum depth measurement of the soil (~1.5 – 2.0 m), and both the 1.0- and 0.5-m dipole data were collected. Measurements were recorded on the device at a rate of

1.0 measurement per second, providing a measurement every 1 m for every walking transect. The EM38 was walked along alternating rows to ensure meter-scale coverage of the field. To minimize drift caused by the effect of temperature, ECa readings were taken as expediently as possible (within 1.0 – 1.5 hours) while maintaining a consistent distance between the EM38 and the ground. The method used to collect the soil EC_a data for all locations was to i) walk the perimeter of the field and then ii) walk alternating rows and then iii) finalizing the scan by walking an 'x' pattern across the entire trial area to collect a set of overlapping data points. The x pattern at the end was a final check for any instrument drift.

Using histograms and a basic classified x-y posting of the data, raw EC_a data was assessed for normality, drift caused by temperature changes, and any extreme outliers. None of the locations used for this study showed evidence of drift, except for Danevang, in which drift appeared in the EC_a ratings of the final 'x' pattern described previously. To address this, the EC_a measurements were plotted with time, the only data that showed drift was the "x" patterned data so it was removed from further analyses of Danevang. To achieve a distribution that was closer to normality, the EC_a data were transformed using a natural log. Empirical variograms of transformed data were fit in R using the *variog* function (Ribeiro and Diggle,; Pebesma, 2004; R Development Core Team, 2011). These variogram models were used to perform ordinary kriging in SAGA GIS software and produce interpolated EC_a grids at 0.5 m (Webster and Oliver, 2007; Conrad et al., 2015). A plot map composed of the boundaries of the field plots created and overlaid onto the ECa grid using QGIS software (QGIS Team, 2018). The Zonal

Statistics plugin in QGIS was then used to extract the mean EC_a from each plot. The mean EC_a values for each plot were used in statistical analyses, as well as the plot grain yield and plant height measurements.

3.3.3 Data Analysis and Statistics

Agronomic data (plant height, grain yield) and EC_a data were checked for outliers and normality using JMP Pro Software (SAS Institute Inc., 1989 – 2017). After outliers were identified and removed using the Huber test (K = 4), we determined no need for transformation for normality using a Shapiro-Wilk test in JMP Pro. The plant data were left untransformed for further analysis. All linear models were created in R using the *lm* function. Three different linear models were used to determine which best explained the variation in each trial location. The first linear model was

$$Y = \alpha_i + \varepsilon,$$

where Y is grain yield or plant height; α is genotype (i); and ε is error. This model is referred to as the G (genotype only) model. The purpose of this model was primarily to test the level of spatial autocorrelation in the residuals and variability present in each trial without using any other terms to partition that variability. The second model is

$$Y = \alpha_i + \beta_j + \varepsilon,$$

where β is the replication or block (j). This model is referred to as the G + Rep model and is the traditional model that plant breeders use to address spatial variability in their trials. The third model is

$$Y = \alpha_i + \gamma_\iota + \delta_k + \varepsilon,$$

where γ is EC_a (ι); and δ is genotype * EC_a (interaction effect, $\dot{k}$). This model is referred to as the G + EC_a model and is an attempt to use interpolated EC_a grids, instead of replicated blocks, to account for spatial variability. Analyses of variance were calculated for the three linear models using the *anova* function in R software.

Each of the six trials was tested for spatial autocorrelation using a Moran's I test, which was calculated using the *lm.morantest* function in R software (Bivand et al., 2013; Bivand and Piras, 2015). The spatial neighborhood for the Moran test was defined using a "rook" neighborhood structure; and binomial weights. A "queen" neighborhood structure was also tested but made no change in the results.

3.4 Results and Discussion

3.4.1 Variograms, Interpolated EC_a Field Maps, and Trial Parameters

Each performance trial in this study was characterized for several parameters of agronomic importance (Table 5). Though 2017 was a relatively wet year and many locations in Texas experienced increased rainfall during the growth period, there was a still a marked difference in precipitation between the driest environment, Gregory (38.5 cm), and the wettest environment, Greenville (105.9 cm).

As expected for performance trials composed of advanced sorghum hybrids, grain yields were relatively high in each location, although there were exceptions that are worth mentioning. The first distinction is that the Perryton trial was exposed to hail on two separate occasions and, as a result, the overall vigor and resulting yields were lower than would be expected. Secondly, the Plainview trial experienced very low yields (Table 5). The cause for this is likely attributable to the impact of bird consumption of

the grain while it was still on the sorghum panicles, which is a common issue for sorghum producers worldwide (Manikowski and Camara-Smeets, 1979). Bird damage in this trial was severe enough to reduce the grain ~95% in some plots. While there was variation for plant height in the performance trials, the differences between the tallest and shortest plots were not very extreme. The reason for the lack of significant differences in plant height is because the performance trials were composed entirely of advanced sorghum hybrids, which are limited in their range of acceptable plant heights due to producer requirements and expectations. Nonetheless, variation for plant height is still an important consideration and the trait is often shown to be correlated to, but not solely predictive of, grain yield in sorghum (Cassady, 1965). Thus, it was considered appropriate to include a statistical analysis of plant height in the present study.

Table 5. Summary of Trial Locations. Summary of six sorghum performance trials throughout Texas in 2017. Information includes precipitation during the growing season, the moisture regime, the minimum, maximum, and mean grain yield and plant height for each location, the coefficient of variation, or C.V. for grain yield and plant height, the aerial trial dimensions, the soil classification for each location, and the minimum and maximum of the plot-level soil EC_a values.

			Yield (kg ha)$^{-1}$				Height (cm)			
Location	**Precipitation (cm)**	**Moisture Regime**	**Min**	**Max**	**Mean**	**CV (%)**	**Min**	**Max**	**Mean**	**CV (%)**
Danevang	61.5	Rainfed	6495	9291	8291	8.8	109	150	131	3.3
Greenville	105.9	Rainfed	7308	9161	8326	7.6	117	157	133	2.8
Gregory	38.5	Rainfed	3791	5819	5049	9.3	94	124	113	4.3
Mexia	77.7	Rainfed	3782	6549	5398	15.8	112	150	129	4.1
Perryton	67.0	Irrigated	6223	9282	7793	10.2	94	145	121	3.5
Plainview	79.2	Irrigated	743	6805	4923	17.8	91	152	116	6.4

				EC_a (mS m)$^{-1}$	
Location	**Trial Length (m)**	**Trial Width (m)**	**Soil Classification**	**Min**	**Max**
Danevang	73.2	30.5	Fine, smectitic, hyperthermic Typic Hapluderts	42	60
Greenville	73.2	22.9	Fine, smectitic, thermic Udic Haplusterts	151	170
Gregory	73.2	22.9	Fine, mixed, superactive, hyperthermic Vertic Calciustolls	142	197
Mexia	73.2	30.5	Fine, smectitic, thermic, Chromic Udic Haplusterts	172	201
Perryton	73.2	30.5	Fine, mixed, superactive, mesic Torrertic Paleustolls	101	116
Plainview	73.2	40.5	Fine, mixed, superactive, thermic Torrertic Paleustolls	57	78

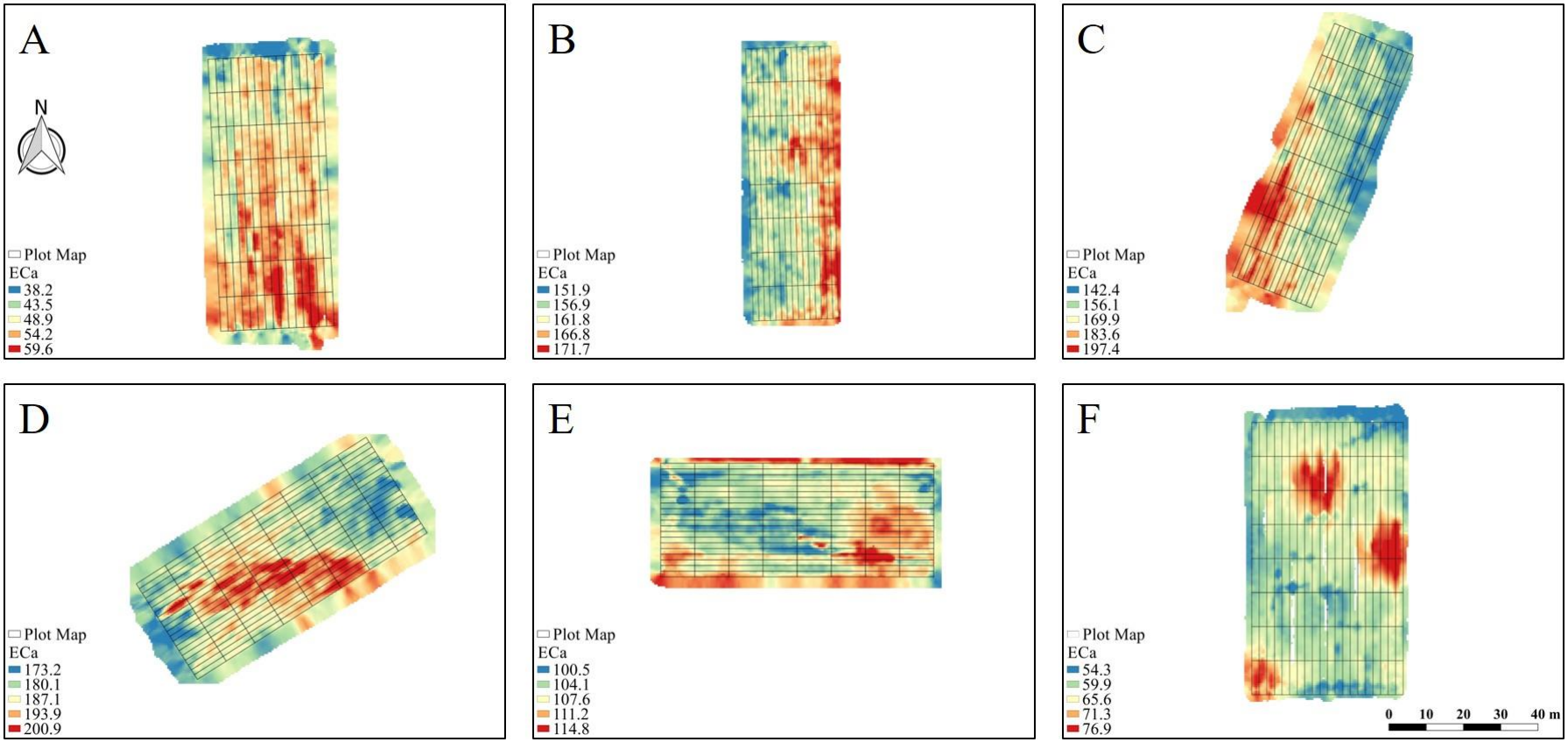
Variogram models of the EC_a data were chosen for each trial location based upon their best fit to the empirical variogram using the coefficient of determination (Table 6). All the variogram models chosen were exponential or spherical and had similar nuggets, ranges, and sills apart from Danevang, which had a very small range (2.3 m) compared to the other five trials (22.9 to 33.1 m). The nuggets were all near zero (0.02 – 0.04 mS m^{-1}).

Each of the interpolated soil EC_a maps exhibited spatial patterns (Figure 5). Interestingly, the patterns of variability in one location, Danevang, seemingly matched the layout of the replicated blocks (Figure 5). At Danvang, the blocks corresponded to the pattern of EC_a values, whereas the first block corresponded to the largest ECA area and the third block corresponded to the smallest-valued ECa area (Figure 5). Conversely, the EC_a patterns were not at all related to the layout of the replicated blocks. For example, in Greenville the areas of highest and lowest EC_a were running throughout all four blocks (Figure 5). Presumably, a more efficient method to block this location may have been to acquire the soil EC_a information prior to planting and plan the trial design in such a way as to address this variation. However, this is conjecture and further experiments, in which trials are laid out in the same location but are blocked in different orientations in subsequent years, would need to be performed to assess the effectiveness of this proposed approach. Other trials of note were the Mexia trial, where a large area of high EC_a cut across the trial diagonally, and the Plainview trial, where several large nodes of high EC_a were present in the field (Figure 5).

Table 6. Summary of Variogram Models. Variograms fit to empirical variogram data, for six sorghum performance trials located throughout Texas in 2017.

Location	Nugget (mS m)$^{-1}$	Range (m)	Sill (mS m)$^{-1}$	Model
Danevang	0.04	2.3	0.09	Exponential
Greenville	0.02	22.9	0.04	Exponential
Gregory	0.03	33.1	0.09	Spherical
Mexia	0.02	23.6	0.04	Spherical
Perryton	0.03	24.7	0.03	Spherical
Plainview	0.03	26.9	0.09	Spherical

Figure 5. Interpolated ECa Maps. Interpolated soil ECa grids for six sorghum performance trials throughout Texas in 2017. The grid of EC_a (mS m)$^{-1}$ values is overlaid with the plot map (black rectangles). Locations include Danevang (A), Greenville (B), Gregory (C), Mexia (D), Perryton (E), and Plainview (F).



3.4.2 Analysis of Variance

For grain yield, analyses of variance demonstrated that, in most cases, the three models did not vary in their model fit, or R^2 values (Table 7). However, there were a few instances where there was a clear improvement of the model from blocking or using soil EC_a. In Greenville, blocking ($R^2 = 0.39$) had a distinct advantage over soil EC_a ($R^2 = 0.28$) as a method for addressing spatial variation. The other example was in Mexia, where the G + EC_a model performed considerably better ($R^2 = 0.63$) than blocking ($R^2 = 0.46$) and was the clear best model for that environment. Other than these two distinctions, the other four trials did not exhibit many differences between the three models (Table 7). While the adjusted R^2 estimate did often increase, the increase was usually quite minimal and was likely artificially inflated due to adding more terms into the associated linear model (Table 7). However, as expected, the error term did decrease when using the G + Rep and G + EC_a models because of some of the error variation being partitioned into the block or EC_a effects, respectively (Table 7). Since bird damage was so prevalent in the Plainview trial, a model including bird damage as an interaction term was conducted, and the resulting analysis of variance did consider the bird damage to be a significant factor in that trial. Thus, the three models presented here were unlikely to account for the variability present (Table 7).

Table 7. Analysis of Variance for Grain Yield. Sums of squares (Sum Sq.) from analysis of variances for plot yield (kg ha)$^{-1}$ using linear models, taken from sorghum performance trials grown in six locations throughout Texas in the year 2017.

		Danevang		**Greenville**		**Gregory**		**Mexia**		**Perryton**		**Plainview**
G	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**
Genotype	29	50073033***	26	26253844**	23	18580675***	35	137903759***	36	116371556**	36	283584690***
Error	80	44551978	80	37947312	72	15934695	104	111015018	109	171392901	95	77882106
Adj. R^2		0.36		0.22		0.39		0.40		0.21		0.70
G + Rep	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**
Genotype	29	50073033***	26	26253844***	23	18580675***	35	137903759***	36	116371556**	36	283584690***
Rep	3	3205722	3	9548974***	3	1102760	3	13611295**	3	5884894	3	7891784*
Error	77	41346255	77	28398338	69	14831935	101	97403723	106	165508006	92	69990322
Adj. R^2		0.38		0.39		0.41		0.46		0.21		0.72
G + EC_a	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**
Genotype	29	50073033***	26	26253844**	23	18580675***	35	137903759***	36	116371556**	36	283584690***
EC_a	1	34209	1	20584	1	3583416***	1	32606733***	1	26196	1	239674
G*EC_a	29	20901795	26	14828461	23	2775205	35	33151515	36	62012343	36	33449407
Error	50	23615973	53	23098267	48	9576074	68	45256770	72	109354361	58	44193026
Adj. R^2		0.46		0.28		0.45		0.63		0.23		0.72

* Significant at $p < 0.05$
** Significant at $p < 0.01$
***Significant at $p < 0.0001$

For plant height, there was also not a very noticeable difference between the three models in most trial locations, although there were a few exceptions again (Table 8). One example was in Gregory (Table 8). While that location showed limited improvement when using the G + Rep or G + EC_a models for grain yield, it does appear that it was significantly improved by using either of these models, although the G + Rep model was superior since the Rep effect was also highly significant, indicating that it was a term that could remain in the model for further analysis (Table 8). Just as was seen in the grain yield analysis of variance, Mexia was improved using either method, although the G + EC_a model was still superior (Table 8). Interestingly, Perryton exhibited some improvement when using the G + Rep model, even though the variability of the plant heights in that location was not expected to be very large due to the hail-induced damage present there (Table 8).

Table 8. Analysis of Variance for Plant Height. Sums of squares (Sum Sq.) from analysis of variance for plant height (cm) using linear models, taken from sorghum performance trials grown in six locations throughout Texas in the year 2017.

		Danevang		**Greenville**		**Gregory**		**Mexia**		**Perryton**		**Plainview**
G	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**
Genotype	29	6668.5***	26	8648.5***	23	2939.5	35	4812.3***	36	14264.7***	36	12378.8***
Error	80	1104.8	81	1295.8	72	2122.2	69	2569.0	111	3030.6	95	5297.8
Adj. R^2		0.81		0.83		0.45		0.48		0.77		0.59
G + Rep	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**
Genotype	29	6668.5***	26	8648.5***	23	2939.5***	35	4812.3***	36	14264.7***	36	12378.8***
Rep	3	26.0	3	156.8*	3	552.7***	2	421.0**	3	1112.2***	3	250.1
Error	77	1078.8	78	1139.0	69	1569.4	67	2148.0	108	1918.4	92	5047.7
Adj. R^2		0.80		0.84		0.57		0.55		0.85		0.59
G + EC_a	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**	**df**	**Sum Sq.**
Genotype	29	6668.5***	26	8648.5***	23	2939.5***	35	4812.3***	36	14264.7***	36	12378.8***
EC_a	1	0.0	1	45.3	1	783.9	1	264.5**	1	69.3	1	392.7*
G*EC_a	29	350.3	26	383.5	23	425.1	35	1336.1	36	607.8	36	1529.8
Error	50	754.5	54	867.0	48	913.2	33	968.4	74	2353.5	58	3375.4
Adj. R^2		0.79		0.83		0.64		0.59		0.73		0.57

* Significant at $p < 0.05$
** Significant at $p < 0.01$
***Significant at $p < 0.0001$

3.4.3 Analysis of Spatial Autocorrelation Using Moran's I

The results of the Moran's I testing for each trial largely corroborated the findings of the analysis of variance. Moran's I test was only concerned with determining the presence or absence of spatial autocorrelation in the residuals (Table 9). For grain yield, the residuals of the G model did not exhibit any spatial autocorrelation for Gregory and Danevang; thus, the G + Rep and G + EC_a models were unnecessary in those instances (Table 9). Greenville did have a significant level of spatial autocorrelation in the residuals. Blocking was successful in making the residuals spatially independent (Table 9). The fact that blocking removed the spatial autocorrelation and the EC_a data did not, indicates a factor other than soil variability was the source of the spatial autocorrelation (Stroup et al., 1994). For example, competition between neighboring cultivars in a trial can be a significant contributing factor in the overall performance of a given genotype (Jensen and Federer, 1964; Kempton and Lockwood, 1984; Stroup et al., 1994). Various areas of the field may also have been subjected to pests such as insects or diseases at varying levels (Sharma, 1993). These few examples help illustrate why the G + Rep model may have been effective while the G + EC_a was not in that environment (Table 9). Mexia exhibited the largest amount of spatial autocorrelation for grain yield, and neither the G + Rep and G + EC_a models were effective at removing it (Table 9). However, in this case the G + EC_a model did perform better than the G + Rep model. In this situation, it is reasonable to conclude that soil variability was more of factor in Mexia than Greenville (Table 9). Perryton exhibited a very low amount of spatial autocorrelation in the residuals, bordering on insignificant, and only the G + EC_a model

was effective in removing it (Table 9). However, as previously stated, it is difficult to assess these factors in that environment since the plants were subjected to hail on two separate occasions. Finally, Plainview exhibited significant amounts of spatial autocorrelation, but neither of the two models was effective at removing it, and this was because of the excessive bird damage observed in that trial (Table 9). After conducting a Moran's I test while including bird damage as an interaction effect, it was confirmed that adding that term to the basic G model removed the spatial autocorrelation present in that location.

Analysis of plant height, provided similar results except for some differences in the Gregory, Mexia, and Perryton trials (Table 9). Interestingly, in Mexia, there was no spatial autocorrelation in the residuals, though grain yield model did have autocorrelation (Table 9). There are a few possible explanations for this. First, as previously discussed, the plants in sorghum performance trials composed of advanced hybrids were unlikely to vary significantly for plant height. Similarly, Gregory may have had spatial autocorrelation for plant heights, but that does not necessarily imply that there should have been spatial autocorrelation for grain yield. Indeed, plant height and grain yield are correlated in sorghum and other grains, but plant height has not been shown to be a reliable sole predictor of final yield performance.

Table 9. Summary of Moran's I Results. Summary of presence or absence of spatial autocorrelation in model residuals of the three models used in the experiment. Autocorrelation is determined by the Moran's I value and its associated test for significance. Locations include six sorghum performance trials throughout Texas in the year 2017.

Grain Yield (kg ha)$^{-1}$				
Location	**G**	**G + Rep**	**G + EC_a**	**Comments**
Danevang	No	No	No	No spatial autocorrelation present
Greenville	Yes (*)	No	Yes (*)	Blocking removed the spatial autocorrelation
Gregory	No	No	No	No spatial autocorrelation present
Mexia	Yes	Yes	Yes (**)	Both models were insufficient to account for the spatial autocorrelation
Perryton	Yes (*)	Yes (*)	No	Very low, almost insignificant amounts of spatial autocorrelation present,
Plainview	Yes (**)	Yes (*)	Yes (**)	Both models were insufficient to account for the spatial autocorrelation
Plant Height (cm)				
Location	**G**	**G + Rep**	**G + EC_a**	**Comments**
Danevang	No	No	No	No spatial autocorrelation present
Greenville	Yes (*)	No	Yes (*)	Blocking removed the spatial autocorrelation
Gregory	Yes (**)	No	No	Both models were sufficient to account for the spatial autocorrelation
Mexia	No	No	No	No spatial autocorrelation present
Perryton	Yes	Yes (**)	Yes	Both models were insufficient to account for the spatial autocorrelation
Plainview	No	No	No	No spatial autocorrelation present

* Significant at $p < 0.05$
** Significant at $p < 0.01$
***Significant at $p < 0.0001$

3.5 Conclusions

Based on these results, it may sometimes be unnecessary to use blocking or soil EC_a maps to address spatial variability in plant breeding trials, but there are certain cases where these methods are required. Thus, it is not realistically viable for plant breeders to forgo blocking their material entirely. Additionally, the cost to the plant breeder to use blocking is minimal, so using RCB or other trial designs that minimize spatial variation is always appropriate. Nonetheless, there are situations where blocking may not be enough to remove spatial variability, and in some of these situations producing soil EC_a grids stands as an attractive option. Other than the one-time cost of the EM38 instrument, the only cost associated with the production of an EC_a map is about 2 hours for the time it takes to walk the area of the trial. Therefore, our results suggest that plant breeders interested in addressing soil variability that may cause spatial autocorrelation of their model residuals should consider using EC_a grids as well as blocking designs and use either on in a post-hoc test of independent residuals in the model. We also expect that in years when rainfall causes stress, soil spatial patterns may form more strongly in resulting variety trials because of variability in soil-stored available water. Further research in other production environments will be required to assess this approach for other agronomic characteristics and in other crops as well as to begin implementing it into modern plant breeding programs.

4. ESTIMATION OF PLANT HEALTH IN A SORGHUM FIELD INFECTED WITH ANTHRACNOSE USING A FIXED-WING UNMANNED AERIAL SYSTEM*

4.1 Synopsis

Diseases cause enormous losses of yield and quality for crop producers worldwide. To meet future food demands, crops are bred for resistance to as many of these maladies as possible. One such disease, anthracnose [*Colletotrichum sublineola*] is a fungal disease of great importance to sorghum [*Sorghum bicolor*, L. Moench] production because it causes significant annual economic losses in the crop. Breeding for anthracnose resistance requires time-consuming phenotyping which is subjective and conditional to the evaluator. It is possible that quantitative assessment using high-throughput methodologies to estimate the trait may be more effective. In this study, we present a statistical analysis of fixed-wing UAS evaluation of anthracnose incidence and severity in sorghum using normalized difference vegetation index. In early phases of infection, correlations between ground-truth and UAS estimates of anthracnose are moderate but they increase to very high by the end of the season (r = -0.55 to -0.95). Additionally, both metrics have moderate to high repeatabilities throughout the growth period (r= 0.60 to 0.90), indicating they are consistently able to differentiate genotypes. Finally, we find that the UAS-derived measurements (R^2 = 0.377, 0.473) are better

* Reprinted with permission from "Estimation of plant health in a sorghum field infected with anthracnose using a fixed-wing unmanned aerial system" by Pugh, N. A., X. Han, S. D. Collins, J. A. Thomasson, D. Cope, A. Chang, J. Jung, T. S. Isakeit, L. K. Prom, G. Carvalho, I. T. Gates, A. Vree, G. C. Bagnall, and W. L. Rooney., 2018, Journal of Crop Improvement, 32:6 681-877, Copyright 2018 by Name of Informa UK Limited, trading as Taylor and Francis Group.

associated with ground-truth measurements (R^2 = 0.278, 0.347) for grain yield under anthracnose pressure. The results of this study indicate that fixed-wing UAS can potentially be effective for evaluating anthracnose disease presence in sorghum, and the greater range of the UAS allows the effective evaluation of larger numbers than ground truth or traditional remote sensing methods.

4.2 Introduction

To meet the food production expectations, cereal crop yields must increase at a rate of at least 2.4% per year (Godfray et al., 2010; Cairns et al., 2013;). Unfortunately, numerous crop diseases reduce yield potentials throughout the world (Strange and Scott, 2005). Phenotyping for disease has historically been difficult and has served as a barrier to the improvement of crop resistance (Furbank and Tester, 2011). High-throughput techniques, particularly those that utilize remote sensing, could reduce this disease phenotyping bottleneck which would help crop improvement programs achieve the required rate of genetic gains for future production (Tester and Langridge, 2010; Furbank and Tester, 2011; Araus and Cairns, 2014; Shi et al., 2016). The capability of unmanned aerial systems (UAS) to cover large areas in a relatively short length of time makes them an appealing option for plant breeders (Shi et al., 2016; Chang et al., 2017; Malambo et al., 2018; Pugh et al., 2018; Shafian et al., 2018). However, no previous studies have conducted an in-depth evaluation of fixed-wing UAS for their ability to measure large numbers of disease breeding plots and make decisions based upon the results.

Previous studies have shown the efficacy of remote sensing to estimate other biotic and abiotic stressors in crops, and these same concepts could apply to the phenotyping of crop diseases (Valasek et al., 2016; Caturegli et al., 2016; Yang et al., 2018). A common remote sensing technique to assess plant stress is normalized difference vegetation index, (NDVI), which is a rating that is often used as a measurement of plant health and photosynthetic integrity (Rouse et al., 1974; Jones and Vaughan, 2010). NDVI has been previously evaluated and implemented as an estimator of many parameters of interest including disease presence, photosynthetic activity, abiotic stress, and others (Rouse et al., 1974; Gamon et al., 1995; Anyamba and Tucker, 2005; Kumar et al., 2016).

Disease progress curves have been utilized by crop scientists for many years and have been applied in breeding programs previously (Néya and Le Normand, 1998; Hess et al., 2002; Li and TeBeest, 2009). The area under the disease progress curve, or AUDPC, can provide a quantitative assessment of disease severity within each individual genotype (Jeger and Viljanen-Rollinson, 2001). In much the same way that individual assessments of disease severity could be useful, disease progress curves generated using high-throughput techniques could produce new information for breeders.

Sorghum (*Sorghum bicolor*, L. Moench) serves as an excellent case study by which to evaluate UAS for their ability to phenotype disease. While there are several diseases of significance in sorghum, anthracnose (*Colletotrichum sublineola* P. Henn., Kabát and Bubák) is arguably the most important (Tebeest et al., 2004; Li and TeBeest,

2009; Moore et al., 2010). In sorghum, the disease infects all above ground plants of the part (stalk, leaves, peduncle, and panicle) producing visible damage (Warren, 1986; Rodriguez-Herrera et al., 2000; Moore et al., 2010). Anthracnose can cause sorghum yield reductions as high as 50% or more (Ali et al., 1987; Li and TeBeest, 2009). As such, anthracnose serves as an excellent case study to assess UASs for their ability to phenotype the same.

Genetic resistance is the primary means of controlling the disease. Because there are many different strains of *C. sublineola*, sorghum improvement scientists must continuously breed for resistance in their material (Leslie, 2002). Traditional methods of phenotyping sorghum for the presence and severity of the disease in the field are ultimately effective but are also quite laborious and are prone to a significant amount of measurer error due to their inherent subjectivity of the rating. Thus, the objectives of this study were to i) evaluate the relationship between remotely-sensed NDVI data taken using a fixed-wing UAS and ground-truth anthracnose disease rating, and to ii) evaluate each of the various measurements in this study for their relationship to final grain yield.

4.3 Materials and Methods

4.3.1 Germplasm and Experimental Design

The germplasm used for the evaluation of anthracnose was composed of a set of 557 experimental sorghum grain hybrids. Within this material, a set of 15 hybrids that were replicated two to seven times throughout the trial to account for spatial variability within the field. These replicated hybrids were used for the calculation of least square Means, variance components, and repeatability. Each hybrid was planted in a 6.71-m test

plot with 1.22-m alleys between them. The entire set of 557 plots was used for correlative analysis. The experimental test was planted in College Station, TX on March 31st, 2017. Standard agronomic practices for grain sorghum were used in this study with the exception that plants were inoculated with *C. sublineola* to ensure disease infection.

4.3.2 Preparation and Application of Anthracnose Inoculant

To prepare anthracnose inoculum, the methods described by Prom et al. (2009) were used. First, nine single-spored isolates of *C. sublineola* were inoculated onto ½ X potato dextrose agar plates with streptomycin (0.1 mg/ml) and were incubated for two weeks at ambient room temperature (~25° C) (Guthrie, 1992). One hundred agar plates were prepared for each of these isolates. After incubation, spores were removed from the plates by placing approximately five mL of water on them and gently dislodging them with a glass rod. Approximately 600 mL were collected in this manner per isolate, and each product was brought to 1000 mL by adding distilled water. This spore suspension was used to inoculate autoclaved sorghum seeds.

The autoclaved seeds were prepared as follows: 3000 cc volume of sorghum seeds and 1.5 L of distilled water were placed into a 46 by 38 by 13 cm autoclavable plastic tray and were covered with aluminum foil. This mixture was autoclaved for one hour, with rapid exhaust (i.e. one minute on the dry cycle). After autoclaving, the tray was emptied into a 132-L translucent plastic bag and the bag was spread across the working area to increase its surface area and increase the rate of cooling. After cooling, 250 mL of inoculum was added to each bag and was shaken thoroughly to mix. Bags of inoculum were incubated open and spread out in a warm room (~28° C) out of direct

sunlight for three to four days prior to application in the field. Each bag was inoculated with one isolate, but prior to field inoculation, inoculum of different isolates was uniformly mixed.

Inoculation of sorghum plants in the test plots occurred 49 days after planting (DAP). To ensure even distribution of the inoculum, about 2 – 3 plants were inoculated per meter by directly sprinkling inoculum into the whorl of the plants. During the growing season, anthracnose developed on the inoculated plants and spores arising from these infections then spread throughout the trial via splash dispersal, a process whereby a fungal pathogen spreads via water droplets (Madden, 2009).

4.3.3 Field Measurements of Disease Incidence and Grain Yield

Ground-truth measurements of anthracnose incidence and severity in the sorghum plots was conducted via a subjective visual rating within the research plot (Table 10). Disease presence was determined depending upon the proportion of the plants in each plot that exhibited any anthracnose symptoms i.e. necrotic lesions, grey diseased tissue, etc. These measurements were recorded on five different dates during the growing season (Table 11). Since the sorghum plots were inoculated with anthracnose at the beginning of the growing season, it was reasonably expected that most of the diseased tissue within those plots was specifically due to anthracnose. Grain yield was measured on the genotypes that were replicated throughout the trial by hand harvesting the plot and threshing it in an Almaco plot thresher.

Table 10. Field Measurements of Anthracnose Incidence and Severity. This table summarizes the 1 – 9 visual rating (Rating) that was used to estimate disease incidence and severity (Description) in sorghum research plots in College Station, TX in 2017 (Pugh et al., 2018).

Rating	Description
1	Disease inconspicuous or present on an occasional plant, only speckling occurs
2	Disease is present and has up to 50% prevalence on all plants with low severity on each plant; apparently causing little damage
3	Disease is present and over 50% prevalence on all plants with low severity on each plant; apparently causing little damage
4	Disease is present with 100% prevalence on all plants; some severity which apparently is causing some damage
5	Disease is severe with 100% prevalence on all plants; estimated leaf area destroyed up to 25%; disease appears to be of economic importance
6	Disease is severe with 100% prevalence on all plants; estimated leaf area destroyed is between 25 to 50%; disease is of economic importance
7	Disease is severe with 100% prevalence on all plants; estimated leaf area destroyed is between 50 to 75%; disease is of economic importance
8	Disease is severe with 100% prevalence on all plants; estimated leaf area destroyed is above 75%; disease is of economic importance
9	Disease is severe with 100% prevalence on all plants; leaf area destroyed is 100%; death of leaves or plants due to disease

Table 11. Summary of Flight and Ground-truth Dates. The dates for the ground-truth (GT) data taken during this study as well as when the NDVI was taken via an unmanned aerial system (NDVI Date) (Pugh et al., 2018). The ground-truth date is also shown as the number of days after planting (DAP), and this value will be used for the rest of this study to refer to the associated GT and UAS dates. In addition, the growth stage of the sorghum as described by Vanderlip and Reeves (1972).

GT Date	GT Date (DAP)	NDVI Date	Sorghum Growth Stage
June 14th, 2017	75	June 16th, 2017	Half-bloom
June 21st, 2017	82	June 23rd, 2017	Soft Dough
July 3rd, 2017	94	June 29th, 2017	Hard Dough
July 18th, 2017	109	July 13th, 2017	Physiological Maturity
July 27th, 2017	118	July 25th, 2017	Physiological Maturity

4.3.4 UAS Data Collection and Data Processing

UAS flights were conducted 24 times from March to August in 2017. We used five datasets collected on June 16th, June 23rd, June 29th, July 13th, and July 25th collected using a Tuffwing ® UAV Mapper fixed-wing platform outfitted with RGB and multispectral camera. The RGB camera used was a Sony ® A6000 with 16 mm Pancake Lens. A RedEdge multispectral sensor, manufactured by MicaSense (Washington, USA), was mounted to collect 5 bands (Red, Green, Blue, Red-edge, Near-infrared) spectral imagery. Immediately after collecting this data, the raw imagery was uploaded to a university data-sharing hub (UASHub) and was made available for processing. UAS images were processed via Agisoft Photoscan Pro software (AgiSoft LLC, St. Petersburg, Russia) with Structure from Motion (SfM) to generate the Digital Surface

Model (DSM) and orthomosaic image (Chang et al., 2017). For precise geo-referencing of the orthomosaic images, 16 Ground Control Points (GCPs) were installed in study area and the coordinates of these GCPs were measured by RTK GPS. The coordinates were input to Agisoft Photoscan Pro when orthomosiac and DSM were generated. The positions of the GCPs were manually elected on one image and automatically projected to the remaining images in ESRI ArcGIS 10.3.1 software. The spatial resolution of the orthomosaic image was 7~8 cm. All orthomosaic images were geo-referenced with less than 4 cm accuracy, which meant a pixel size of 0.5.

The mosaicked images for the experimental field were then clipped with the 597 research plots within it, which formed a region of interest (ROI) and 4.7 m^2 (6.0 by 0.78 m) for each. To prevent plot edge effects from influencing calculations, the plots were reduced by 0.2 m on each edge using an ArcGIS buffer tool. Data from the multispectral cameras were converted into normalized difference vegetation indices (NDVI) to estimate the incidence of anthracnose disease. The NDVI was calculated based upon the equation

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}}$$

where ρ_{NIR} and ρ_{Red} are the reflectance of the near-infrared and the visible red bands, respectively.

4.3.5 Data Analysis and Statistics

The area under the disease progress curve (AUDPC) is an informative metric that can be used to temporally estimate the severity of disease in crops (Jeger and Velamen-

Rollinson, 2000). The AUDPC for anthracnose incidence within the experimental material was calculated in Microsoft Excel software using the equation:

$$AUDPC = \sum_{i=1}^{n} (\frac{y_i + y_{i+1}}{2})(t_{i+1} - t_i)$$

, where n = total number of observations, y_i = disease intensity or incidence at the ith observation, and t = time at the ith observation. The AUDPC metric was calculated using the date of inoculation, and areas under the disease curves were calculated for each ground-truth date.

Sorghum disease estimates and AUDPCs collected and calculated via ground-truth and via the UAS were checked for outliers using the Huber test in JMP Pro 12.2.0 software (SAS Institute Inc., 1989 - 2017). Least squares means were calculated for the ground-truth measurements and the UAS-derived NDVI estimates in JMP by using an analysis of variance (ANOVA, all fixed). The same model was used for a restricted maximum likelihood analysis (REML) that was conducted within environments using Fit Model (all random) in JMP. The statistical model that was used to perform these analyses was,

$$Y = \alpha_i + \gamma_\iota + \delta_k + \varepsilon$$

where Y is NDVI or ground-truth estimates of anthracnose; α is genotype (i); γ is the row index (ι); δ is the range (k); and ε is the error. For the REML results, effects that had negative variance components were subsequently removed from that model. The percentage of total genotypic variation as well as the repeatability (R) estimates were calculated using this model. Repeatability estimates were calculated using the equation:

$$R = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$$

where σ_g^2 = the genotypic variance, and σ_e^2 = the error variance (Nakagawa and Schielzeth, 2010). The least squares means, REML, and repeatability estimates were obtained using the replicated hybrids within the trial.

Variance components and repeatability estimates were obtained for NDVI and the daily visual rating scores. Using all plots, Pearson's correlation coefficients (*r*) were calculated between ground-truth estimates of each parameter and the UAS-derived estimates (SAS Institute Inc., 1989 - 2017). In addition, simple linear regression was performed between the four estimates of disease and grain yield using the replicated hybrids in JMP. The yield analysis was conducted to better understand which measurement is most useful to predict yield losses due to anthracnose.

4.4 Results and Discussion

4.4.1 Anthracnose Infection and Development

The general progression of anthracnose throughout the trial can be clearly observed via the visual rating (Ground-truth) and NDVI-derived estimates (Figure 6, Figure 7). Overall, anthracnose incidence and severity increased as the growth period continued, though the increase was much more pronounced in the NDVI data (Figure 7). The reason for this is not entirely known, but it is likely due to the inherent nature of NDVI. An NDVI rating is a vegetation index that provides an assessment of overall plant health and is not necessarily specific only to anthracnose presence. Thus, NDVI

values started quite high on the first flight date (75 DAP) but fell markedly by the end of the season (118 DAP).

Figure 6. Least Squares Means for Ground-truth Estimates of Anthracnose Disease. This figure shows the general progression of anthracnose estimates via visual rating (Ground-truth) (Pugh et al., 2018). Estimates are presented as least squares means (L. S. Means) across the five measurement dates, which are presented as the days after planting (DAP).

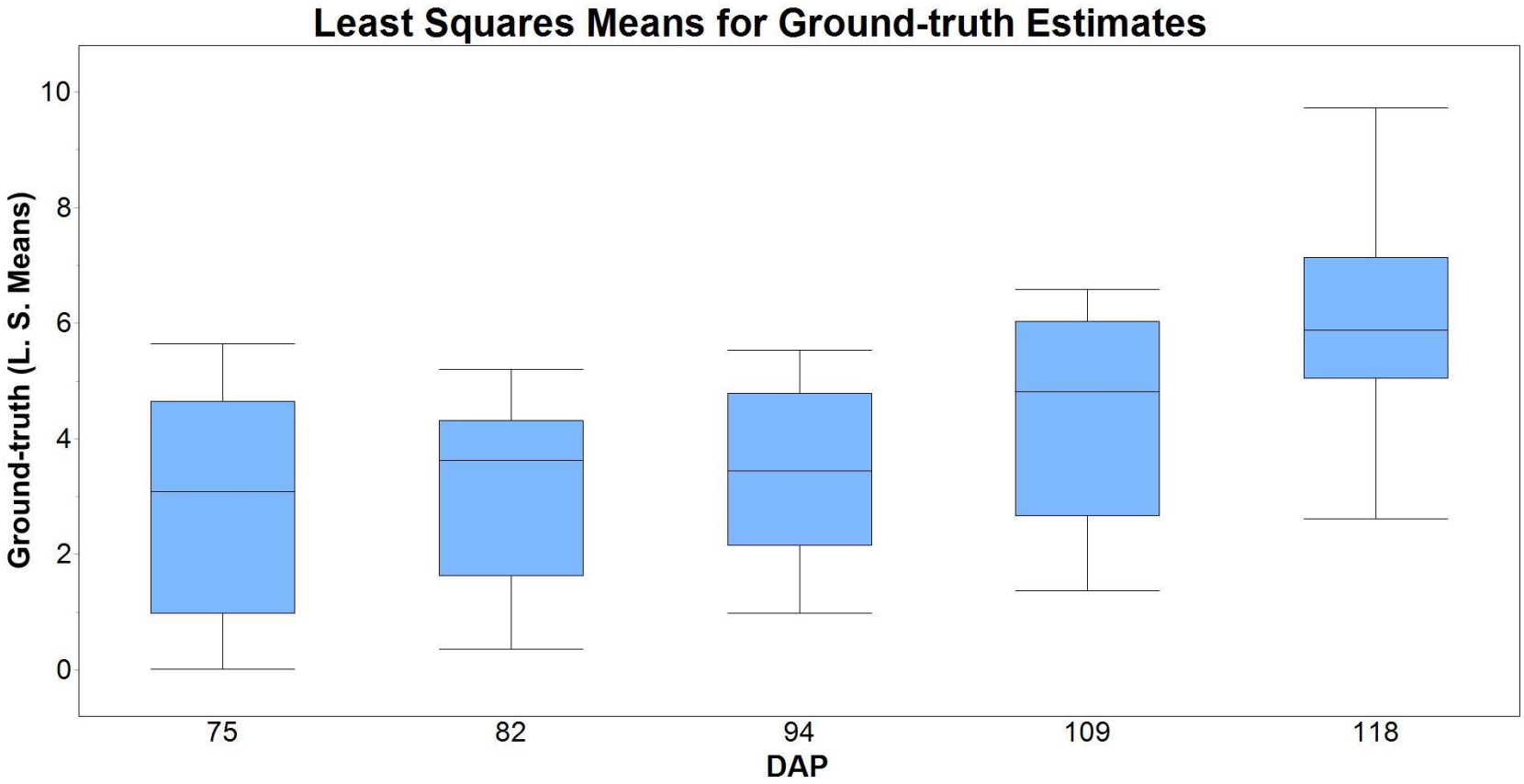
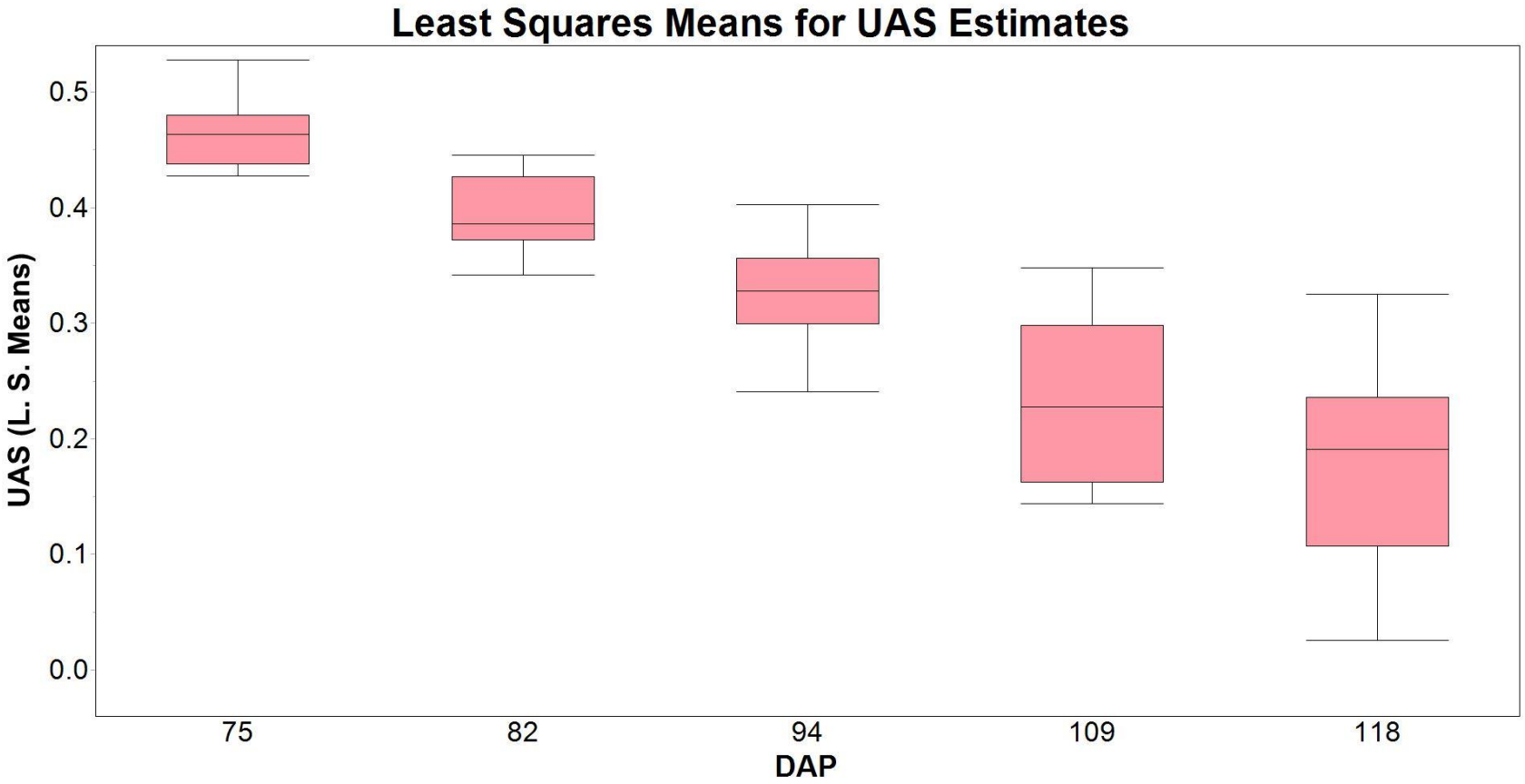


Figure 7. Least Squares Means for NDVI Estimates of Anthracnose Disease. This figure shows the general progression of anthracnose estimates via normalized difference vegetation index, or NDVI (Pugh et al., 2018). Estimates are presented as least squares means (L. S. Means) across the five measurement dates, which are presented as the days after planting (DAP).



4.4.2 Correlations Between Ground-truth and UAS Estimates of Anthracnose Incidence and Severity in Sorghum

Pearson's correlations (*r*) between ground-truth and NDVI measurements generally increased during the growth period (Table 11). When comparing ground-truth measurements with the UAS-derived NDVI measurement, the correlations were low at the beginning of the season and rapidly increased until reaching a maximum of > -0.90 by the end of growth (Table 11). Similarly, the relationship between the AUDPC generated by the NDVI data (AUDPC-NDVI) and the two ground-truth measurements became stronger over the course of the growth period; however, the correlation was not as strong between those metrics.

Table 12. Correlations for Sorghum Disease Estimates. Pearson's correlation coefficients (*r*) between measurements of disease obtained from either ground-truth or a UAS (Pugh et al., 2018). The ground-truth measures include visual ratings of disease incidence within the plots (Ground-truth) as well as the area under the disease progress curve (AUDPC-GT). The normalized difference vegetation index, or NDVI, and the area under the disease progress curve using NDVI data (AUDPC-NDVI) were derived from the UAS platform. Areas under the disease progress curves were calculated using the equation: $\boldsymbol{AUDPC} = \sum_{i=1}^{n}(\frac{y_i + y_{i+1}}{2})(\boldsymbol{t_{i+1}} - \boldsymbol{t_i})$, n = total number of observations, y_i = disease intensity or incidence at the ith observation, and t = time at the ith observation. Measurements were taken on five flight dates and their closest corresponding ground-truth dates during the growth period, shown as days after planting (DAP). All correlations were statistically significant at the $p < 0.001$ level.

DAP	Ground-truth vs. NDVI	AUDPC-GT vs. NDVI	Ground-truth vs. AUDPC-NDVI	AUDPC-GT vs. AUDPC-NDVI
75	-0.53	-0.53	-0.53	-0.53
82	-0.69	-0.69	-0.58	-0.58
94	-0.74	-0.74	-0.63	-0.65
109	-0.83	-0.81	-0.72	-0.74
118	-0.94	-0.91	-0.80	-0.82

The correlations in this study were negative because resistant plots would have higher NDVI values and AUDPC-NDVIs; conversely, susceptible plots would have higher ground-truth values and AUDPC-GTs (Table 11) (Gamon et al., 1995; Jones and Vaughan, 2010). Because anthracnose was inoculated early (approximately at panicle initiation), most of the foliar damage to the plants that was observed was caused by the disease as opposed to foliar damage by other diseases (Gamon et al., 1995; Jones and Vaughan, 2010). In addition, the presence of anthracnose was repeatedly confirmed on the plants by researchers on the ground. Evaluation dates later in the season (later DAPs) probably have a stronger relationship as the disease is more obvious and the range of variation between plots is greater. This expectation is borne out in the progression of

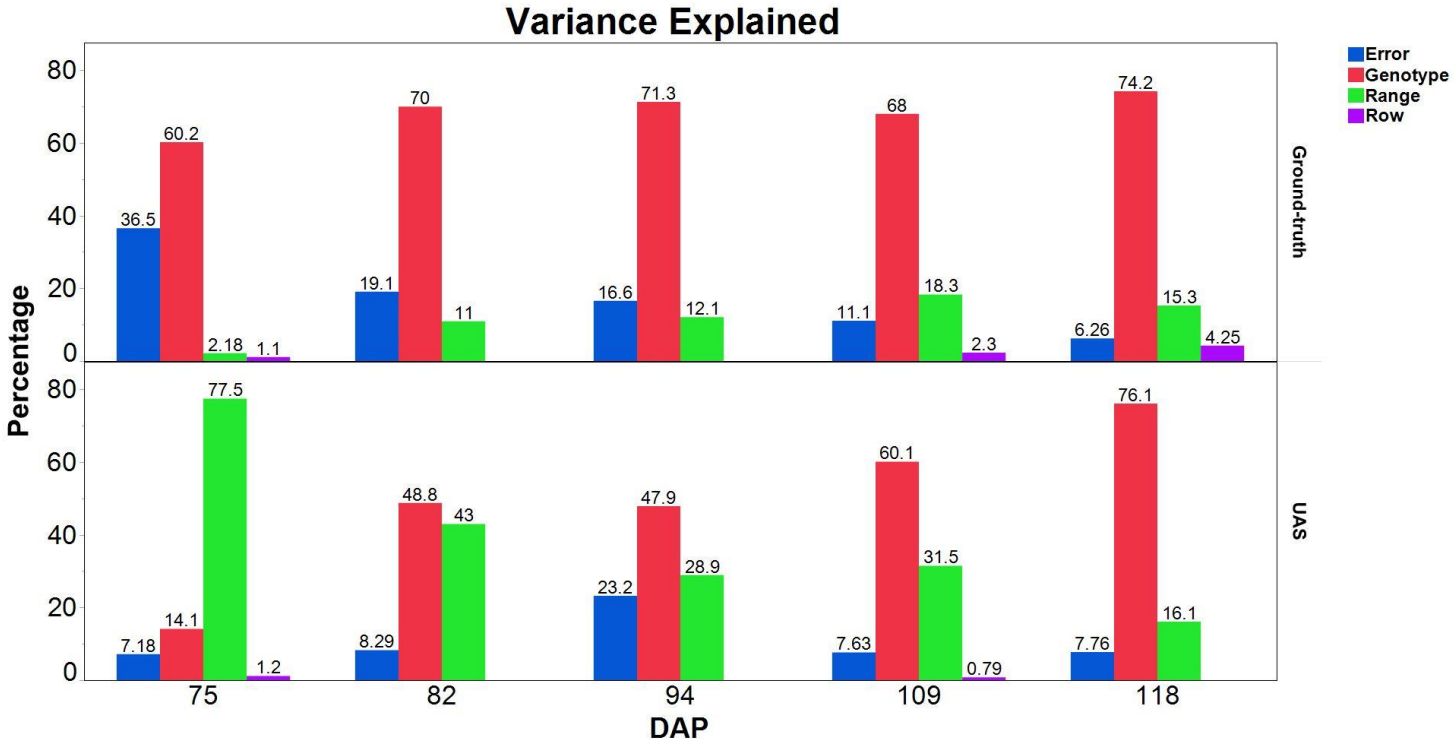
correlation coefficients, as the correlations at the beginning of the season between all ground-truth methods and their respective NDVI measurements were low at < -0.55, and then progressively improved to end season values that ranged in the -0.80's and -0.90's, depending upon the traits being correlated (Table 11). Additionally, upon examination of the raw imagery for each flight date, it was much easier to identify susceptible plots on later flight dates (109 and 118 DAP) than it was for the earlier dates (75 and 82 DAP).

Several prior studies used NDVI estimates to estimate disease presence in crops (Franke and Menz, 2007; Kumar et al., 2016; Pretorius et al., 2017). In Franke and Menz (2007), NDVI results were used to classify the disease severity of research plots infected with powdery mildew (*Blumeria graminis*) and leaf rust (*Puccinia recondita*) in wheat (*Triticum aestivum*). The first date in that study had the lowest overall accuracy at 56.8%, and gradually increased to 88.6% by the last date (Franke and Menz, 2007). Pretorius et al. (2017) and Kumar et al. (2016) both associated NDVI and disease incidence by identifying quantitative trait loci (QTL) that were co-identified using ground-truth estimates. In Kumar et al. (2016), a negative correlation coefficient (-0.91) was observed between the spot blotch (*Bipolaris sorokiniana*) severity measured by field researchers and the NDVI measured at the same growth stage in wheat. Additionally, a QTL for spot blotch resistance was identified in the same interval using the NDVI and ground-truth measures of disease severity (Kumar et al., 2016). Similarly, strong relationships between NDVI and final wheat leaf stripe rust (*Puccinia striiformis*) severity were observed in Pretorius et al. (2017), and four QTL were identified for the trait as well as for NDVI, two of which matched.

4.4.3 Genotypic Variance Explained and Repeatability for Ground-truth and UAS Measurements of Anthracnose Incidence in Sorghum

Ground-truth anthracnose estimates could discern significant genotypic variance between plots from the first date onward (Figure 8). The genotypic variance explained increased over the course of the growth period; additionally, the error variance abruptly decreased from the first to the second date (75 and 82 DAP, respectively) and further decreases were more gradual (Figure 8). In contrast, the UAS-derived NDVI measurements of genotypic variation were minimal but increased as the plants matured (Figure 8).

Figure 8. Variance Explained for Ground-truth and NDVI Estimates of Disease. The differences in the amount of variance that were explained by visual scores (Ground-truth) and normalized difference vegetation index, or NDVI estimates of disease presence in hybrid sorghum grown in College Station, TX in 2017 (Pugh et al., 2018). The statistical model used for this analysis was $Y = \alpha_i + \gamma_\iota + \delta_k + \varepsilon$, where α = genotype (i), γ = row (ι), δ = range (k), and ε = error. Measurements were taken on five flight dates and their closest corresponding ground-truth dates during the growth period, shown as days after planting (DAP).



The ground-truth method was superior for early estimates of disease incidence and severity in the sorghum research plots (Figure 8) (Dudley and Moll, 1969). However, it was matched by the UAS later in the season, wherein the differences between plots were much easier for the NDVI to discern (Figure 8). This indicates that researchers who wish to accurately assess the disease presence within their plots using a fixed-wing UAS need to wait for full development of the disease. The genotypic variance for the ground-truth data was not temporally linear; it varied between measurement dates (Figure 8). These fluctuations within the ground truth data were likely attributable to human error.

Spatial variation was also detected in the analysis. Interestingly, a very high amount of the variance for the NDVI data recorded at 75 DAP was explained by the Range effect; one possible explanation for this is that it could be due to an inability for the multispectral sensor to measure the leaves that were lower in the canopy within each plot, since the typical progression of anthracnose disease begins at the bottom of the sorghum plant and moves upwards (Figure 8). Thus, while researchers on the ground may have been able to spot lesions and other signs of foliar damage on the lower leaves of a sorghum plant early on, remote sensing techniques could not. The technique also cannot differentiate between anthracnose susceptible and resistance responses. In the host pathosystem, some resistant lines may have necrosis or reddening of the leaves without the presence of acervuli, or fungal fruiting bodies, and the presence of acervuli is what indicates the successful reproduction of the pathogen. Another possible explanation for the high Range effect is that the NDVI was measuring differences in the

field that were outside the purview of this study, including nutrient availability, water holding capacity, insect pressure, and others (Figure 8). The high Range effect decreased on subsequent dates, perhaps as the spatial differences within the field began to be outweighed by differences in susceptibility between plots (Figure 8). Further experiments will be necessary to fully determine which of these explanations, if any, is correct.

For ground-truth repeatability, the estimates were moderate early but quickly increased to high as the season progressed (Table 13). A similar trend was observed in the NDVI data, with exception for the 94 DAP flight where the repeatability dropped to a value that was closer to the estimate at 75 DAP (Table 13). Since repeatability is calculated from the genotypic and error variances, it is likely that the score was reduced due to how close these two values were to each other on that flight date (Figure 8) (Nakagawa and Schielzeth, 2010). The reason for the increased error and the reduced genotypic variance on that date is unknown, but possible reasons include extenuating weather and flight conditions as well as image overlap issues. Nonetheless, UAS based NDVI repeatabilities recovered and repeatabilities for the remainder of the trial were similar regardless of whether the NDVI or the visual rating was used (109 and 118 DAP) (Table 13). These results demonstrate that the UAS is as consistent as the ground-truth methodology later in the season but is less consistent during early stages of growth.

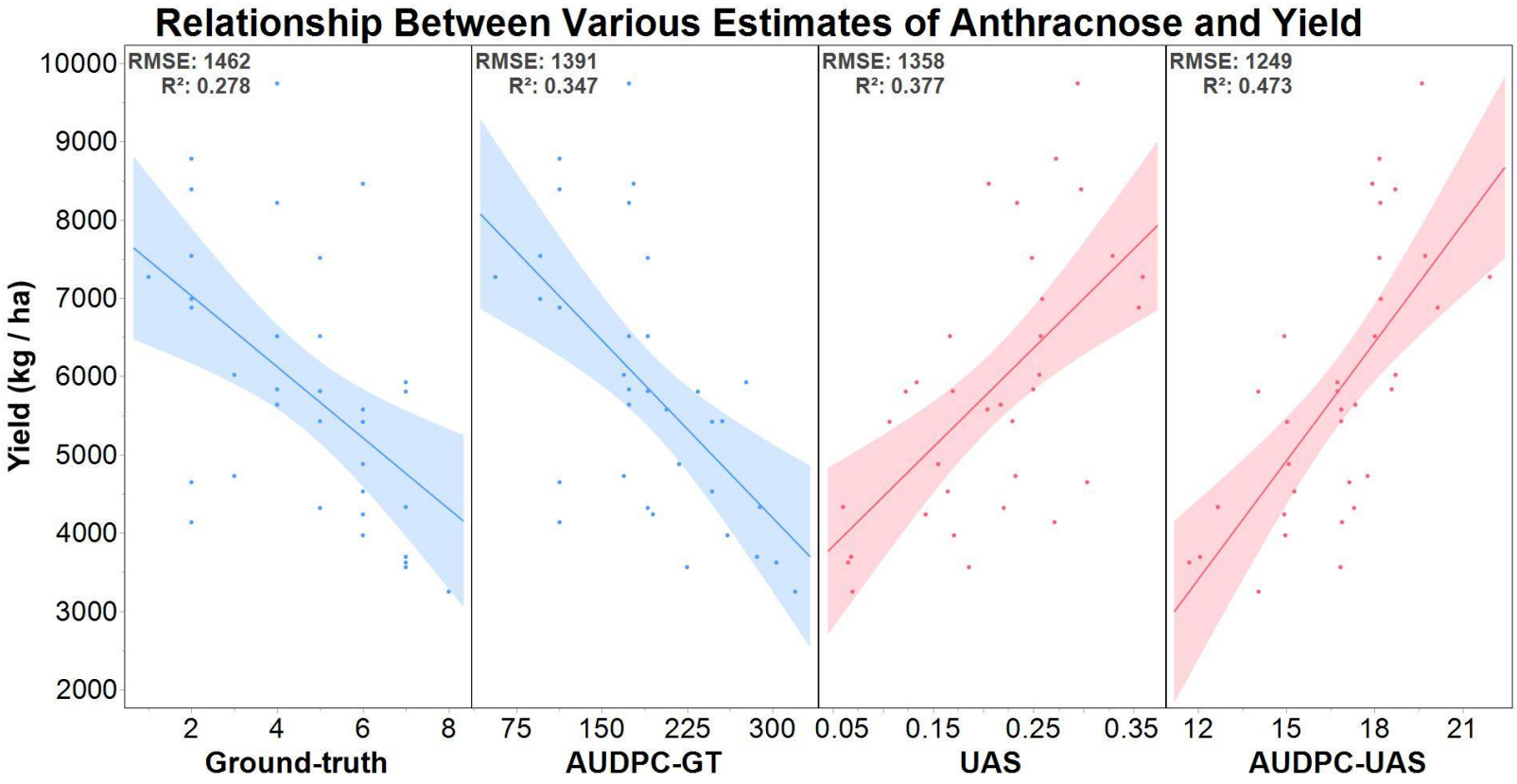
Table 13. Repeatability Estimates. The repeatability (*R*) estimates for a visual rating (Ground-truth) and normalized difference vegetation index, or NDVI, calculated over the course of five flight dates and their closest corresponding ground-truth dates in sorghum (Pugh et al., 2018). The dates where measurements were taken are shown as days after planting (DAP). Estimates were calculated using the equation, $R = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{}}$, where *R* = the repeatability score, σ_g^2 = the genotypic variance, and σ_e^2 = the error variance.

DAP	Ground-truth	NDVI
75	0.62	0.66
82	0.77	0.86
94	0.81	0.67
109	0.86	0.89
118	0.92	0.91

4.4.4 Ground-truth and UAS-derived Measurements of Anthracnose and Their Relationship with Grain Yield

Based upon simple linear regressions, UAS measurements using NDVI were superior for predicting final yield losses because of disease; indeed, the ground-truth methods (Ground-truth) lagged far behind their counterparts in this capacity (Figure 9). The daily scores were less effective than the AUDPCs that were generated from them, the individual scores were also statistically inferior and were subject to more error as they had higher root mean square error (RMSE) values (Figure 9). Thus, plant breeders that wish to predict the loss of yield that can be attributed to disease within their plots are perhaps best served by using the AUDPC-NDVI.

Figure 9. Regression of Disease Measurements and Yield. The relationship between various measurements of disease presence and severity and final grain yield (kg/ha) in sorghum (Pugh et al., 2018). Ground-truth measurements include the subjective visual rating (Ground-truth) that was taken daily, as well as the area under the disease progress curve generated using the same data (AUDPC-GT). Unmanned aerial system (UAS) measurements include normalized difference vegetation index (NDVI) and the AUDPC generated using NDVI (AUDPC-NDVI).



Previous studies have determined that there is a strong relationship between disease severity and grain yield (Gaunt, 1995; Savary et al., 2000). However, no studies have evaluated fixed-wing UAS for their ability to predict yield losses due to disease. The results of this study suggest that the AUDPC-NDVI measurement is a better indicator of disease severity over the course of the entire season since it has a stronger relationship with yield; however, it is important to note that this phenomenon could also be attributed to the fact that NDVI is a measure of overall plant health as previously mentioned (Rouse et al., 1974; Thomas et al., 1996; Li and TeBeest, 2009; Jones and Vaughan, 2010). Therefore, it is not surprising that it is a more useful predictor of yield than measurements that only estimate the presence of disease. Further study will be required to determine the exact relationship between NDVI and final grain yield in sorghum in the presence of anthracnose infection.

4.5 Conclusions

This study has been the first statistical evaluation of a fixed-wing UAS for its ability to estimate anthracnose disease incidence and severity in sorghum. Based on the results herein, unmanned aerial systems can potentially serve as estimators of anthracnose incidence and severity in sorghum, so long as the measurements are taken later in the growth period. Conversely, ground-truth estimates taken using a subjective visual rating are superior during early periods of growth. In addition, ground-truth methodologies can explain a higher amount of genotypic variation during early stages of growth but are quickly matched by NDVI. However, the relationship between the UAS-derived NDVI measurements and final grain yield is stronger than the correlation

between yield and ground-truth disease estimates. Thus, researchers can reliably use NDVI and AUDPCs generated using NDVI as predictors of yield loss due to anthracnose, provided that those measurements are used at the end of growth. Of course, this can also be attributed to the fact that NDVI is measuring differences between genotypes that may not be due to anthracnose itself, since that measurement is an estimate of overall plant health. Future studies should elucidate more information about this relationship and determine how and when these technologies should be implemented in disease breeding programs. Nonetheless, this study serves as a proof-of-concept that fixed-wing UAS can be used to estimate anthracnose disease presence in sorghum breeding programs.

5. CONCLUSIONS

In this dissertation, evidence of the effectiveness of proximal and remote sensing technologies in a sorghum breeding program has been presented. Remote sensing via rotary-winged UAS are effective at phenotyping plant height and biomass yield in sorghum using the VOL and P95 measurements, respectively. In addition, we have shown the potential utility of multitemporal growth curves for analyzing these traits. The growth curve technique could serve as a useful tool for plant breeders in the future, and further studies will be required to elucidate what new determinations can be drawn for it as well as its application within a breeding program. We have also shown how proximal sensing can be used to estimate the EC_a of the soil in several sorghum performance trials. Though most trials did not require blocking or EC_a grids to account for spatial autocorrelation, since four out of the six trials did not show spatial autocorrelation to begin with, both techniques had at least one environment where they were superior to the other. Therefore, we propose that plant breeders interested in accounting for spatial variability in their trials consider the use of both techniques. Then, the proper model can be determined by testing for the presence or absence of spatial autocorrelation via Moran's I tests. Finally, the effectiveness of using NDVI estimates derived from fixed-wing UAS imagery to phenotype anthracnose disease incidence and severity was evaluated. We found that NDVI is an effective estimator of anthracnose presence later in the growing season but is less effective than ground-truth methods early on, presumably because anthracnose spreads from the bottom of the sorghum plants upward and is covered by the canopy early on. Thus, the diseased plants are not visible to UAS-

mounted sensors in early stages of growth. Nonetheless, NDVI serves as a potential tool that sorghum breeders can use to make selections for anthracnose resistance so long as the estimates are taken at the end of the growing season.

To summarize, we have shown that proximal and remote sensing technologies are successful at accomplishing various important tasks within a breeding program. Though the techniques worked well for their various applications in this study, it is important to note that further evaluation of these technologies is required. Not only can their uses described herein be expanded upon in other crops or other environments, but there are plenty of other characteristics of sorghum that have yet to be measured by these techniques. Potentially, new phenotypes and methods of accounting for variation may arise from these studies. Proximal and remote sensing, and by extension high-throughput phenotyping, serves as an important paradigm shift for breeders and will help breeders to meet future food and fiber demands in a changing world.

REFERENCES

Ali, M., H. Warren, R.L.-P. Disease, and undefined 1987. Relationship between anthracnose leaf blight and losses in grain yield of sorghum. cabdirect.org. https://www.cabdirect.org/cabdirect/abstract/19881106253 (accessed 6 September 2018).

Anderson-Cook, C., … M.A.-… S.S. of, and U. 2002. Differentiating soil types using electromagnetic conductivity and crop yield maps. dl.sciencesocieties.org. https://dl.sciencesocieties.org/publications/sssaj/abstracts/66/5/1562 (accessed 6 September 2018).

Anderson-Cook, C.M., M.M. Alley, J.K.F. Roygard, R. Khosla, R.B. Noble, and J.A. Doolittle. 2002. Differentiating Soil Types Using Electromagnetic Conductivity and Crop Yield Maps. Soil Sci. Soc. Am. J. 66(5): 1562. doi: 10.2136/sssaj2002.1562.

Anyamba, A., and C.J. Tucker. 2005. Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. J. Arid Environ. 63(3): 596–614. doi: 10.1016/J.JARIDENV.2005.03.007.

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19(1): 52–61. doi: 10.1016/J.TPLANTS.2013.09.008.

Bendig, J., K. Yu, H. Aasen, A. Bolten, S. Bennertz, J. Broscheit, M.L. Gnyp, and G. Bareth. 2015. International Journal of Applied Earth Observation and Geoinformation Combining UAV-based plant height from crop surface models , visible , and near infrared vegetation indices for biomass monitoring in barley. Int.

J. Appl. Earth Obs. Geoinf. 39: 79–87. doi: 10.1016/j.jag.2015.02.012.

Cairns, J.E., J. Hellin, K. Sonder, J.L. Araus, J.F. MacRobert, C. Thierfelder, and B.M. Prasanna. 2013. Adapting maize production to climate change in sub-Saharan Africa. Food Secur. 5(3): 345–360. doi: 10.1007/s12571-013-0256-x.

Calviño, M. 2012. Sweet sorghum as a model system for bioenergy crops. Curr. Opin. Biotechnol. 23(3): 323–329. doi: 10.1016/J.COPBIO.2011.12.002.

Cassady, A.J. 1965. Effect of a Single Height (Dw) Gene of Sorghum on Grain Yield, Grain Yield Components, and Test Weight1. Crop Sci. 5(5): 385. doi: 10.2135/cropsci1965.0011183X000500050002x.

Caturegli, L., M. Corniglia, M. Gaetani, N. Grossi, S. Magni, M. Migliazzi, L. Angelini, M. Mazzoncini, N. Silvestri, M. Fontanelli, M. Raffaelli, A. Peruzzi, and M. Volterrani. 2016. Unmanned Aerial Vehicle to Estimate Nitrogen Status of Turfgrasses (X Hu, Ed.). PLoS One 11(6): e0158268. doi: 10.1371/journal.pone.0158268.

Chang, A., J. Jung, M.M. Maeda, and J. Landivar. 2017. Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). Comput. Electron. Agric. 141: 232–237. doi: 10.1016/J.COMPAG.2017.07.008.

Chapman, S., T. Merz, A. Chan, P. Jackway, S. Hrabar, M. Dreccer, E. Holland, B. Zheng, T. Ling, J. Jimenez-Berni, S.C. Chapman, T. Merz, A. Chan, P. Jackway, S. Hrabar, M.F. Dreccer, E. Holland, B. Zheng, T.J. Ling, and J. Jimenez-Berni. 2014. Pheno-Copter: A Low-Altitude, Autonomous Remote-Sensing Robotic Helicopter for High-Throughput Field-Based Phenotyping. Agronomy 4(2): 279–301. doi:

10.3390/agronomy4020279.

Colomina, I., and P. Molina. 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS J. Photogramm. Remote Sens. 92: 79–97. doi: 10.1016/J.ISPRSJPRS.2014.02.013.

Conrad, O., B. Bechtel, M. Bock, H. Dietrich, E. Fischer, L. Gerlitz, J. Wehberg, V. Wichmann, and J. Böhner. 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8(7): 1991–2007. doi: 10.5194/gmd-8-1991-2015.

Cooper, M., F. Technow, C. Messina, C. Gho, and L.R. Totir. 2016. Use of Crop Growth Models with Whole-Genome Prediction: Application to a Maize Multienvironment Trial. Crop Sci. 56(5): 2141. doi: 10.2135/cropsci2015.08.0512.

Corwin, D.L., and E. Scudiero. 2016. Field-Scale Apparent Soil Electrical Conductivity. Methods Soil Anal. 1(1): 0. doi: 10.2136/methods-soil.2015.0038.

Dandois, J.P., E.C. Ellis, J.P. Dandois, and E.C. Ellis. 2010. Remote Sensing of Vegetation Structure Using Computer Vision. Remote Sens. 2(4): 1157–1176. doi: 10.3390/rs2041157.

Fehr, W.R. 1987. Principles of Cultivar Development. Macmillan.

Franke, J., and G. Menz. 2007. Multi-temporal wheat disease detection by multi-spectral remote sensing. Precis. Agric. 8(3): 161–172. doi: 10.1007/s11119-007-9036-y.

Furbank, R.T., and M. Tester. 2011. Phenomics – technologies to relieve the phenotyping bottleneck. Trends Plant Sci. 16(12): 635–644. doi: 10.1016/J.TPLANTS.2011.09.005.

Gamon, J.A., C.B. Field, M.L. Goulden, K.L. Griffin, A.E. Hartley, G. Joel, J. Penuelas, and R. Valentini. 1995. Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. Ecol. Appl. 5(1): 28–41. doi: 10.2307/1942049.

Gaunt, R.E. 1995. The Relationship Between Plant Disease Severity and Yield. Annu. Rev. Phytopathol. 33(1): 119–144. doi: 10.1146/annurev.py.33.090195.001003.

Gill, J.R., P.S. Burks, S.A. Staggenborg, G.N. Odvody, R.W. Heiniger, B. Macoon, K.J. Moore, M. Barrett, and W.L. Rooney. 2014. Yield Results and Stability Analysis from the Sorghum Regional Biomass Feedstock Trial. BioEnergy Res. 7(3): 1026–1034. doi: 10.1007/s12155-014-9445-5.

Godfray, H.C.J., J.R. Beddington, I.R. Crute, L. Haddad, D. Lawrence, J.F. Muir, J. Pretty, S. Robinson, S.M. Thomas, and C. Toulmin. 2010. Food security: the challenge of feeding 9 billion people. Science 327(5967): 812–8. doi: 10.1126/science.1185383.

Grenzdörffer, G.J. Crop height determination with UAS point clouds DAISI-Aerial sensor system and technical infrastructure to monitor seabirds and marine mammals View project PFIFFikus-Innovative Photogrammetrie für Micro UAV's View project CROP HEIGHT DETERMINATION WITH UAS POINT CLOUDS. doi: 10.5194/isprsarchives-XL-1-135-2014.

Guthrie, P. A. I., Magill, C. W., Frederiksen, R. A., Odvody, G.N. 1992. Random Amplified Polymorphic DNA Markers: A System for Identifying and Differentiating Isolates of Colletotrichum graminicola. Phytopathology 82(8): 832–

835. https://www.apsnet.org/publications/phytopathology/backissues/Documents/1992Articles/Phyto82n08_832.PDF (accessed 7 September 2018).

Hämmerle, M., and B. Höfle. 2016. Direct derivation of maize plant and crop height from low-cost time-of-flight camera measurements. Plant Methods 12(1): 50. doi: 10.1186/s13007-016-0150-6.

Hess, D.E., R. Bandyopadhyay, and I. Sissoko. 2002. Pattern Analysis of Sorghum Genotype × Environment Interaction for Leaf, Panicle, and Grain Anthracnose in Mali. Plant Dis. 86(12): 1374–1382. doi: 10.1094/PDIS.2002.86.12.1374.

Holman, F.H., A.B. Riche, A. Michalski, M. Castle, M.J. Wooster, and M.J. Hawkesford. 2016. High Throughput Field Phenotyping of Wheat Plant Height and Growth Rate in Field Plot Trials Using UAV Based Remote Sensing. doi: 10.3390/rs8121031.

Hoshmand, R. 2006. Design of Experiments for Agriculture and the Natural Sciences : Design of Experiments for Agriculture and the Natural Sciences. doi: 10.1201/9781420010640.

Iqbal, F., A. Lucieer, K. Barry, R. Wells, F. Iqbal, A. Lucieer, K. Barry, and R. Wells. 2017. Poppy Crop Height and Capsule Volume Estimation from a Single UAS Flight. Remote Sens. 9(7): 647. doi: 10.3390/rs9070647.

Jeger, M.J., and S.L.H. Viljanen-Rollinson. 2001. The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. TAG Theor. Appl. Genet. 102(1): 32–40. doi: 10.1007/s001220051615.

Jensen, N.F., and W.T. Federer. 1964. Adjacent Row Competition in Wheat1. Crop Sci. 4(6): 641. doi: 10.2135/cropsci1964.0011183X000400060027x.

Johnson, C.K., and K.M. Eskridge. 2005. Apparent soil electrical conductivity: applications for designing and evaluating field-scale experiments. Comput. Electron. Agric. 46(1–3): 181–202. doi: 10.1016/J.COMPAG.2004.12.001.

Johnson, C.K., K.M. Eskridge, and D.L. Corwin. 2005. Apparent soil electrical conductivity : applications for designing and evaluating field-scale experiments. 46: 181–202. doi: 10.1016/j.compag.2004.12.001.

Jones, H. G., Vaughan, R.A. 2010. Remote Sensing of Vegetation: Principles, Techniques, and Applications - Hamlyn G Jones, Robin A Vaughan - Google Books. Oxford University Press.

Kempton, R.A., and G. Lockwood. 1984. Inter-plot competition in variety trials of field beans (Vicia faba L.). J. Agric. Sci. 103(02): 293. doi: 10.1017/S0021859600047249.

Kitchen, N.R., S.T. Drummond, E.D. Lund, K.A. Sudduth, and G.W. Buchleiter. 2003. Soil Electrical Conductivity and Topography Related to Yield for Three Contrasting Soil–Crop Systems. Agron. J. 95(3): 483–495. doi: 10.2134/AGRONJ2003.4830.

Kitchen, N.R., K.A. Sudduth, and S.T. Drummond. 1999. Soil Electrical Conductivity as a Crop Productivity Measure for Claypan Soils. jpa 12(4): 607. doi: 10.2134/jpa1999.0607.

Kumar, S., M.S. Röder, R.P. Singh, S. Kumar, R. Chand, A.K. Joshi, and U. Kumar.

2016. Mapping of spot blotch disease resistance using NDVI as a substitute to visual observation in wheat (Triticum aestivum L.). Mol. Breed. 36(7): 95. doi: 10.1007/s11032-016-0515-6.

Leslie, J.F., and Wiley InterScience (Online service). 2002. Sorghum and millets diseases. Iowa State Press.

Li, Y., and D.O. TeBeest. 2009. Temporal and Spatial Development of Sorghum Anthracnose in Arkansas. Plant Dis. 93(3): 287–292. doi: 10.1094/PDIS-93-3-0287.

Machado, S., E.D. Bynum, T.L. Archer, R.J. Lascano, L.T. Wilson, J. Bordovsky, E. Segarra, K. Bronson, D.M. Nesmith, and W. Xu. 2002. Spatial and Temporal Variability of Corn Growth and Grain Yield. Crop Sci. 42(5): 1564. doi: 10.2135/cropsci2002.1564.

Madden, L. V. 2009. Canadian Journal of Plant Pathology Effects of rain on splash dispersal of fungal pathogens. doi: 10.1080/07060669709500557.

Major, D.J., S.B. Rood, and F.R. Miller. 1990. Temperature and Photoperiod Effects Mediated by the Sorghum Maturity Genes. Crop Sci. 30(2): 305. doi: 10.2135/cropsci1990.0011183X003000020012x.

Malambo, L., S.C. Popescu, S.C. Murray, E. Putman, N.A. Pugh, D.W. Horne, G. Richardson, R. Sheridan, W.L. Rooney, R. Avant, M. Vidrine, B. McCutchen, D. Baltensperger, and M. Bishop. 2018. Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. Int. J. Appl. Earth Obs. Geoinf. 64: 31–42. doi: 10.1016/J.JAG.2017.08.014.

Manikowski, S., and M. Da Camara-Smeets. 1979. Estimating Bird Damage to Sorghum and Millet in Chad. J. Wildl. Manage. 43(2): 540. doi: 10.2307/3800369.

Matese, A., S.F. Di Gennaro, and A. Berton. 2017. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. Int. J. Remote Sens. 38(8–10): 2150–2160. doi: 10.1080/01431161.2016.1226002.

McNeil, J.D. 1980. Electrical conductivity of soils and rocks: genoics limited technical note TN-5. Mississauga, Ontario, Canada.

Moore, J.W., M. Ditmore, and D.O. TeBeest. 2010. Development of Anthracnose on Grain Sorghum Hybrids Inoculated with Recently Described Pathotypes of *Colletotrichum sublineolum* Found in Arkansas. Plant Dis. 94(5): 589–595. doi: 10.1094/PDIS-94-5-0589.

Murphy, R.L., R.R. Klein, D.T. Morishige, J.A. Brady, W.L. Rooney, F.R. Miller, D. V Dugas, P.E. Klein, and J.E. Mullet. 2011. Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. Proc. Natl. Acad. Sci. U. S. A. 108(39): 16469–74. doi: 10.1073/pnas.1106212108.

Nakagawa, S., and H. Schielzeth. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol. Rev. 85(4): no-no. doi: 10.1111/j.1469-185X.2010.00141.x.

Neely, H.L., C.L.S. Morgan, C.T. Hallmark, K.J. McInnes, and C.C. Molling. 2016. Apparent electrical conductivity response to spatially variable vertisol properties. Geoderma 263: 168–175. doi: 10.1016/J.GEODERMA.2015.08.040.

Néya, A., and M. Le Normand. 1998. Responses of sorghum genotypes to leaf anthracnose (Colletotrichum graminicola) under field conditions in Burkina Faso. Crop Prot. 17(1): 47–53. doi: 10.1016/S0261-2194(98)80012-4.

Patrignani, A., and T.E. Ochsner. 2015. Canopeo: A Powerful New Tool for Measuring Fractional Green Canopy Cover. Agron. J. 107(6): 2312. doi: 10.2134/agronj15.0150.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30(7): 683–691. doi: 10.1016/J.CAGEO.2004.03.012.

Pretorius, Z.A., C.X. Lan, R. Prins, V. Knight, N.W. McLaren, R.P. Singh, C.M. Bender, and F.J. Kloppers. 2017. Application of remote sensing to identify adult plant resistance loci to stripe rust in two bread wheat mapping populations. Precis. Agric. 18(4): 411–428. doi: 10.1007/s11119-016-9461-x.

Pugh, N.A., D.W. Horne, S.C. Murray, G. Carvalho, L. Malambo, J. Jung, A. Chang, M. Maeda, S. Popescu, T. Chu, M.J. Starek, M.J. Brewer, G. Richardson, and W.L. Rooney. 2017. Temporal Estimates of Crop Growth in Sorghum and Maize Breeding Enabled by Unmanned Aerial Systems. Plant Phenome J. 1(1). doi: 10.2135/TPPJ2017.08.0006.

R Development Core Team, R. 2011. R: A Language and Environment for Statistical Computing.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050 (JP Hart, Ed.). PLoS One 8(6): e66428. doi: 10.1371/journal.pone.0066428.

Ray, D.K., N. Ramankutty, N.D. Mueller, P.C. West, and J.A. Foley. 2012. Recent patterns of crop yield growth and stagnation. Nat. Commun. 3(1): 1293. doi: 10.1038/ncomms2296.

Rhoades, J.D., P.A.C. Raats, and R.J. Prather. 1976. Effects of Liquid-phase Electrical Conductivity, Water Content, and Surface Conductivity on Bulk Soil Electrical Conductivity1. Soil Sci. Soc. Am. J. 40(5): 651. doi: 10.2136/sssaj1976.03615995004000050017x.

Ribeiro, P.J., and P. Diggle. GeoR: A Package for Geostatistical Analysis Compositional data View project The pontine nuclei View project.

Robert, P.C., R.H. Rust, W.E. Larson, D.B. Jaynes, T.S. Colvin, and J. Ambuel. 1995a. Yield Mapping by Electromagnetic Induction. p. 383–394. *In* Site-Specific Management for Agricultural Systems. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.

Robert, P.C., R.H. Rust, W.E. Larson, K.A. Sudduth, N.R. Kitchen, D.F. Hughes, and S.T. Drummond. 1995b. Electromagnetic Induction Sensing as an Indicator of Productivity on Claypan soils. p. 671–681. *In* Site-Specific Management for Agricultural Systems. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.

Rodriguez-Herrera, R., W.L. Rooney, D.T. Rosenow, and R.A. Frederiksen. 2000. Inheritance of grain mold resistance in grain sorghum without a pigmented testa. Crop Sci. 40(6): 1573–1578. doi: 10.2135/cropsci2000.4061573x.

Rooney, W.L., and S. Aydin. 1999. Genetic Control of a Photoperiod-Sensitive

Response in Sorghum bicolor (L.) Moench. Crop Sci. 39(2): 397. doi: 10.2135/cropsci1999.0011183X0039000200016x.

Rooney, W.L., J. Blumenthal, B. Bean, and J.E. Mullet. 2007. Designing sorghum as a dedicated bioenergy feedstock. Biofuels, Bioprod. Biorefining 1(2): 147–157. doi: 10.1002/bbb.15.

Roth, L., and B. Streit. 2018. Predicting cover crop biomass by lightweight UAS-based RGB and NIR photography: an applied photogrammetric approach. Precis. Agric. 19(1): 93–114. doi: 10.1007/s11119-017-9501-1.

Rouse, J.W.. J., R.H. Haas, J.A. Schell, and D.W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS. https://ntrs.nasa.gov/search.jsp?R=19740022614 (accessed 9 September 2018).

Salas Fernandez, M.G., P.W. Becraft, Y. Yin, and T. Lübberstedt. 2009. From dwarves to giants? Plant height manipulation for biomass yield. Trends Plant Sci. 14(8): 454–461. doi: 10.1016/J.TPLANTS.2009.06.005.

Sankaran, S., L.R. Khot, C.Z. Espinoza, S. Jarolmasjed, V.R. Sathuvalli, G.J. Vandemark, P.N. Miklas, A.H. Carter, M.O. Pumphrey, N.R. Knowles, and M.J. Pavek. 2015. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. Eur. J. Agron. 70: 112–123. doi: 10.1016/J.EJA.2015.07.004.

Savary, S., L. Willocquet, F. a. Elazegui, N.P. Castilla, and P.S. Teng. 2000. Rice Pest Constraints in Tropical Asia: Quantification of Yield Losses Due to Rice Pests in a Range of Production Situations. Plant Dis. 84(3): 357–369. doi:

10.1094/PDIS.2000.84.3.357.

Schnell, R., K. Horn, D. Pietsch, S. Hirst, A. Hall, and W.L. Rooney. 2017. Grain Sorghum Performance Trials in Texas Department of Soil and Crop Sciences 2017 GRAIN SORGHUM PERFORMANCE TESTS IN TEXAS.

Shafian, S., N. Rajan, R. Schnell, M. Bagavathiannan, J. Valasek, Y. Shi, and J. Olsenholler. 2018. Unmanned aerial systems-based remote sensing for monitoring sorghum growth and development (JL Gonzalez-Andujar, Ed.). PLoS One 13(5): e0196605. doi: 10.1371/journal.pone.0196605.

Sharma, H.C. 1993. Host-plant resistance to insects in sorghum and its role in integrated pest management. Crop Prot. 12(1): 11–34. doi: 10.1016/0261-2194(93)90015-B.

Shi, Y., J.A. Thomasson, S.C. Murray, N.A. Pugh, W.L. Rooney, S. Shafian, N. Rajan, G. Rouze, C.L.S. Morgan, H.L. Neely, A. Rana, M. V. Bagavathiannan, J. Henrickson, E. Bowden, J. Valasek, J. Olsenholler, M.P. Bishop, R. Sheridan, E.B. Putman, S. Popescu, T. Burks, D. Cope, A. Ibrahim, B.F. McCutchen, D.D. Baltensperger, R. V. Avant, M. Vidrine, and C. Yang. 2016. Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research (J Zhang, Ed.). PLoS One 11(7): e0159781. doi: 10.1371/journal.pone.0159781.

Soil Survey Staff, Natural Resources Conservation Service, U.S.D. of A. Official Soil Series Descriptions. : Available Online.

De Souza, C.H.W., R.A.C. Lamparelli, J.V. Rocha, and P.S.G. Magalhães. 2017. Height estimation of sugarcane using an unmanned aerial system (UAS) based on structure from motion (SfM) point clouds. Int. J. Remote Sens. 38(8–10): 2218–2230. doi:

10.1080/01431161.2017.1285082.

Stroppiana, D., M. Migliazzi, V. Chiarabini, A. Crema, M. Musanti, C. Franchino, and P. Villa. 2015. Rice yield estimation using multispectral data from UAV: A preliminary experiment in northern Italy. p. 4664–4667. *In* 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE.

Stroup, W.W., P.S. Baenziger, and D.K. Mulitze. 1994. Removing Spatial Variation from Wheat Yield Trials: A Comparison of Methods. Crop Sci. 34(1): 62. doi: 10.2135/cropsci1994.0011183X003400010011x.

Team, Q.D. 2018. QGIS Geographic Information System. Open Source Geospatial Found. Proj.

Tebeest, D., T. Kirkpatrick, and R. Cartwright. 2004. Common and Important Diseases of Grain Sorghum. Coop. Ext. Serv. Publ. MP-297: 37–46. https://www.uaex.edu/publications/pdf/mp297/6_diseases.pdf (accessed 9 September 2018).

Tester, M., and P. Langridge. 2010. Breeding technologies to increase crop production in a changing world. Science 327(5967): 818–22. doi: 10.1126/science.1183700.

Valasek, J., J. V. Henrickson, E. Bowden, Y. Shi, C.L.S. Morgan, and H.L. Neely. 2016. Multispectral and DSLR sensors for assessing crop stress in corn and cotton using fixed-wing unmanned air systems. p. 98660L. *In* Valasek, J., Thomasson, J.A. (eds.), International Society for Optics and Photonics.

Walter, A., B. Studer, and R. Kölliker. 2012. Advanced phenotyping offers opportunities for improved breeding of forage and turf species. Ann. Bot. 110(6): 1271–1279.

doi: 10.1093/aob/mcs026.

Warren, H.L. 1986. Leaf anthracnose. Compend. Sorghum Dis. St. Paul, Am. Phytopathol. Soc.: 10–11.

Watanabe, K., W. Guo, K. Arai, H. Takanashi, H. Kajiya-Kanegae, M. Kobayashi, K. Yano, T. Tokunaga, T. Fujiwara, N. Tsutsumi, and H. Iwata. 2017. High-Throughput Phenotyping of Sorghum Plant Height Using an Unmanned Aerial Vehicle and Its Application to Genomic Prediction Modeling. Front. Plant Sci. 8: 421. doi: 10.3389/fpls.2017.00421.

Webster, R. (Richard), and M.A. Oliver. 2007. Geostatistics for environmental scientists. Wiley.

Wheeler, T., and J. von Braun. 2013. Climate change impacts on global food security. Science 341(6145): 508–13. doi: 10.1126/science.1239402.

White, J.W., P. Andrade-Sanchez, M.A. Gore, K.F. Bronson, T.A. Coffelt, M.M. Conley, K.A. Feldmann, A.N. French, J.T. Heun, D.J. Hunsaker, M.A. Jenks, B.A. Kimball, R.L. Roth, R.J. Strand, K.R. Thorp, G.W. Wall, and G. Wang. 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133: 101–112. doi: 10.1016/J.FCR.2012.04.003.

Yang, S., X. Yang, and J. Mo. 2018. The application of unmanned aircraft systems to plant protection in China. Precis. Agric. 19(2): 278–292. doi: 10.1007/s11119-017-9516-7.

Yin, X., M.A. McClure, N. Jaja, D.D. Tyler, and R.M. Hayes. 2011. In-Season Prediction of Corn Yield Using Plant Height under Major Production Systems.

Agron. J. 103(3): 923. doi: 10.2134/agronj2010.0450.

Yue, J., G. Yang, C. Li, Z. Li, Y. Wang, H. Feng, B. Xu, J. Yue, G. Yang, C. Li, Z. Li, Y. Wang, H. Feng, and B. Xu. 2017. Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models. Remote Sens. 9(7): 708. doi: 10.3390/rs9070708.